

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2005, 10:57:20 ; Search time 40.2909 Seconds  
(without alignments)  
3570.253 Million cell updates/sec

Title: US-10-665-283-1  
Perfect score: 9903  
Sequence: 1 MALRGFCSDGSLDWMNV.....RSVAVAKAKRFSIPDLSL 1927

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/iaa/FACTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7860	79.4	1531	1 US-08-463-092B-4	Sequence 4, Appli
2	7860	79.4	1531	2 US-08-462-109A-4	Sequence 4, Appli
3	7860	79.4	1531	2 US-08-460-907B-4	Sequence 4, Appli
4	7860	79.4	1531	3 US-08-463-179A-4	Sequence 4, Appli
5	7860	79.4	1531	3 US-08-461-384B-4	Sequence 4, Appli
6	7860	79.4	1531	4 US-09-647-140B-19	Sequence 19, Appli
7	7849	79.3	1531	1 US-08-141-893-2	Sequence 2, Appli
8	7849	79.3	1531	1 US-08-463-092B-2	Sequence 2, Appli
9	7849	79.3	1531	2 US-08-462-109A-2	Sequence 2, Appli
10	7849	79.3	1531	2 US-08-460-907B-2	Sequence 2, Appli
11	7849	79.3	1531	3 US-08-463-179A-2	Sequence 2, Appli
12	7849	79.3	1531	3 US-08-461-384B-2	Sequence 2, Appli
13	7849	79.3	1531	3 US-08-407-207A-2	Sequence 2, Appli
14	7002.5	70.7	1528	1 US-08-463-092B-6	Sequence 6, Appli
15	7002.5	70.7	1528	2 US-08-462-109A-6	Sequence 6, Appli
16	7002.5	70.7	1528	2 US-08-460-907B-6	Sequence 6, Appli
17	7002.5	70.7	1528	3 US-08-463-179A-6	Sequence 6, Appli
18	7002.5	70.7	1528	3 US-08-461-384B-6	Sequence 6, Appli
19	4487.5	45.3	1527	4 US-09-647-140B-6	Sequence 33, Appli
20	4487.5	45.3	1530	4 US-09-647-140B-33	Sequence 3, Appli
21	3391.5	34.2	1503	4 US-09-647-140B-8	Sequence 8, Appli
22	3385.5	34.2	1503	4 US-09-792-616-3	Sequence 3, Appli
23	3218.5	32.5	1498	4 US-09-792-616-9	Sequence 9, Appli
24	2374.5	24.0	1621	3 US-08-972-927-3	Sequence 3, Appli
25	2349.5	23.7	1622	3 US-08-972-927-6	Sequence 6, Appli
26	2266.5	22.9	1325	4 US-09-647-140B-2	Sequence 2, Appli
27	2195.5	22.2	1361	3 US-09-439-313-538	Sequence 538, App

ALIGNMENTS

RESULT 1  
US-08-463-092B-4  
; Sequence 4, Application US/08463092B  
; Patent No. 5766880  
; GENERAL INFORMATION:  
; APPLICANT: Cole, Susan P.C.  
; APPLICANT: Deeley, Roger G.  
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ENCODING  
; TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS  
; STREET: Queen's University at Kingston  
; CITY: Kingston  
; STATE: Ontario  
; COUNTRY: CANADA  
; ZIP: K7L 3N6  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/463,092B  
; FILING DATE: 05-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/966,923  
; FILING DATE: 27-OCT-1992  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/029,340  
; FILING DATE: 8-MAR-1993  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/141,893  
; FILING DATE: 26-OCT-1993  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/407,207  
; FILING DATE: 20-MAR-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Steeg, Carol Miernicki  
; REGISTRATION NUMBER: 39,539  
; REFERENCE/DOCKET NUMBER: Q1546  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (613) 545-2342  
; TELEFAX: (613) 545-6853  
; INFORMATION FOR SEQ ID NO: 4:

28 2195.5 22.2 1261 4 US-09-636-215-538 Sequence 538, App  
29 2195.5 22.2 1261 4 US-09-685-166A-538 Sequence 538, App  
30 2195.5 22.2 1261 4 US-09-679-426-538 Sequence 538, App  
31 2195.5 22.2 1261 4 US-09-759-143-538 Sequence 538, App  
32 2195.5 22.2 1261 4 US-09-651-236-538 Sequence 538, App  
33 2118 21.4 1581 3 US-08-726-320-3 Sequence 3, Appli  
34 2118 21.4 1581 3 US-09-208-716-3 Sequence 3, Appli  
35 2115 21.4 1228 3 US-09-439-313-537 Sequence 537, App  
36 2115 21.4 1228 4 US-09-636-215-537 Sequence 537, App  
37 2115 21.4 1228 4 US-09-685-166A-537 Sequence 537, App  
38 2115 21.4 1228 4 US-09-679-426-537 Sequence 537, App  
39 2115 21.4 1228 4 US-09-759-143-537 Sequence 537, App  
40 2115 21.4 1228 4 US-09-651-236-537 Sequence 537, App  
41 2082 21.0 1437 3 US-09-061-400-2 Sequence 2, Appli  
42 2082 21.0 1453 2 US-09-001-273-2 Sequence 2, Appli  
43 2082 21.0 1453 3 US-08-843-459A-2 Sequence 38, Appli  
44 2081 21.0 1437 4 US-10-162-012-38 Sequence 4, Appli  
45 2080 21.0 1437 4 US-09-647-140B-4 Sequence 4, Appli

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 1531 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLSCULE TYPE: protein
US-08-463-092B-4

Query Match 79.4%; Score 7860; DB 1; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCAGSDPLWDMNVTWNTSNPDFTKCFONTVLVWPCFYLMACFPFIFYLSRH 60
DB 1 MALRGFCAGSDPLWDMNVTWNTSNPDFTKCFONTVLVWPCFYLMACFPFIFYLSRH 60
QY 61 DRGYIQMTPLNKTALGFLWICWADLFYSFWERSRGIFLAPVFLVSPILLGTTLLA 120
DB 61 DRGYIQMTPLNKTALGFLWICWADLFYSFWERSRGIFLAPVFLVSPILLGTTLLA 120
QY 121 TFLIQERRKGQSSGIMLTFWLVALCALAILRSKIMTALKEDAQVDFRDIIFYVYFS 180
DB 121 TFLIQERRKGQSSGIMLTFWLVALCALAILRSKIMTALKEDAQVDFRDIIFYVYFS 180
QY 181 LLLIQLVLSFSDRSPLFSETHDNPCESSASFLSRITFWITGLIVRGYRQPLEGSD 240
DB 181 LLLIQLVLSFSDRSPLFSETHDNPCESSASFLSRITFWITGLIVRGYRQPLEGSD 240
QY 241 LWSLNKEDTSEQVPLVKNWKECAKTRKOPKVYVSSKOPAKESKVDANEVEAL 300
DB 241 LWSLNKEDTSEQVPLVKNWKECAKTRKOPKVYVSSKOPAKESKVDANEVEAL 300
QY 301 IVKSPQKEWNPFLFKVLYKTRGYPFLMGFFPKAIHDLMMFSGPQILKLLIKFVNDTKAPD 360
DB 301 IVKSPQKEWNPFLFKVLYKTRGYPFLMGFFPKAIHDLMMFSGPQILKLLIKFVNDTKAPD 360
QY 361 WQGYFYTVLLFWTACQLVLHQPFIHFVSGMRKTAIVGAVRKALVIITNSARKSSTV 420
DB 361 WQGYFYTVLLFWTACQLVLHQPFIHFVSGMRKTAIVGAVRKALVIITNSARKSSTV 420
QY 421 GEIVNLSVDAQRFMDLATYINMISAPLOVTLALYLLMLNIGPSVLGAVVLMVPMVN 480
DB 421 GEIVNLSVDAQRFMDLATYINMISAPLOVTLALYLLMLNIGPSVLGAVVLMVPMVN 480
QY 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLYAWELAFKDLAIROBELKVLK 540
DB 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLYAWELAFKDLAIROBELKVLK 540
QY 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTTDENNILDQAQAFVSLALFNILRPLNLP 600
DB 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTTDENNILDQAQAFVSLALFNILRPLNLP 600
QY 601 MVISSIVQASVSLKRLRIFLGSHELEPDSIERRPVKGGGTNSITVRNATFTWARSDDPT 660
DB 601 MVISSIVQASVSLKRLRIFLGSHELEPDSIERRPVKGGGTNSITVRNATFTWARSDDPT 660
QY 661 LINGITFSIPEGALVAVVGQCGKSSLLSALLAEMDKVEGHVAIKGSVAYVPOQAWIQND 720
DB 661 LINGITFSIPEGALVAVVGQCGKSSLLSALLAEMDKVEGHVAIKGSVAYVPOQAWIQND 720
QY 721 SLRENILFGCCLEBPYRSVIOACALLPDLBILPSGDRTEIGEGVNLSSGQKQKVSILAR 780
DB 721 SLRENILFGCCLEBPYRSVIOACALLPDLBILPSGDRTEIGEGVNLSSGQKQKVSILAR 780
QY 781 AVYSNADILYFDDPLSAVDHVGKHIENVIIPGKMLKNKTRILVTHSMSYLPQVDVILV 840
DB 781 AVYSNADILYFDDPLSAVDHVGKHIENVIIPGKMLKNKTRILVTHSMSYLPQVDVILV 840
QY 841 MSGGKISMGYSQELLARDGAFBFLRTYASTEQDAEENGVTGVSQPGKEAKQEMNGM 900
DB 841 MSGGKISMGYSQELLARDGAFBFLRTYASTEQDAEENGVTGVSQPGKEAKQEMNGM 900
QY 901 LVTDGAKQOLQSSSSSYSGDISRHNSAEALQKAEAKKEETWKLMEADKAQTQGVKL 960

```

RESULT 2

```

US-08-462-109A-4
; Sequence 4, Application US/08462109A
; Patent No. 5882875
; GENERAL INFORMATION:
; APPLICANT: Cole, Susan P.C.
; APPLICANT: Deeley, Roger G.
; TITLE OF INVENTION: METHODS FOR IDENTIFYING
; TITLE OF INVENTION: MULTIDRUG RESISTANT TUMOR CELLS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/462,109A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/966,923
; FILING DATE: 27-OCT-1992
; APPLICATION NUMBER: 08/029,340

```

; FILING DATE: 8-MAR-1993
; APPLICATION NUMBER: 08/141,893
; FILING DATE: 26-OCT-1993
; APPLICATION NUMBER: 08/407,207
; FILING DATE: 20-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Decont, Giulio A. Jr.
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: PQI-002CP4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1531 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLSCULE TYPE: protein
; US-08-462-109A-4
Query Match 79.4%; Score 7860; DB 2; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALRGCAGSDPLMDNVTWNTSNPDKFQNTLVVWPCFYLMACFPFYLYLSRH 60
DB 1 MALRGCAGSDPLMDNVTWNTSNPDKFQNTLVVWPCFYLMACFPFYLYLSRH 60
QY 61 DRYGTOMPLNKTALGFLWVWADLFYFWERSRGIPLAPVLSPTLGLITLLA 120
DB 61 DRYGTOMPLNKTALGFLWVWADLFYFWERSRGIPLAPVLSPTLGLITLLA 120
QY 121 TPLQLERRKGVSQSGIMFTLWLVLCALALIRSKIMTALKEDAQVDLFRDITVYVYFS 180
DB 121 TPLQLERRKGVSQSGIMFTLWLVLCALALIRSKIMTALKEDAQVDLFRDITVYVYFS 180
QY 181 LLLIQLVSCFSDRSPLETHDNPCESSASFLSRITFWITGLIVRGVQPLEGSD 240
DB 181 LLLIQLVSCFSDRSPLETHDNPCESSASFLSRITFWITGLIVRGVQPLEGSD 240
QY 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPVVYSSKDPKQKESKVDANEVEAL 300
DB 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPVVYSSKDPKQKESKVDANEVEAL 300
QY 301 IVKSPQKWNPSLFVLYKTFPGPYFLMSPFFKAIHDLMMFSGPQILKLLIKFVNDTKAPD 360
DB 301 IVKSPQKWNPSLFVLYKTFPGPYFLMSPFFKAIHDLMMFSGPQILKLLIKFVNDTKAPD 360
QY 361 WQGYFYTVLLFVTAQLTLVHLQYFHI CFVSGMRITKAVIGAVYRKALVITNSARKSTV 420
DB 361 WQGYFYTVLLFVTAQLTLVHLQYFHI CFVSGMRITKAVIGAVYRKALVITNSARKSTV 420
QY 421 GEIVNLSVDAQRFMDLATYINMIASAPLOVILALYLLWNLGSPSLAGVAVMLVMPVN 480
DB 421 GEIVNLSVDAQRFMDLATYINMIASAPLOVILALYLLWNLGSPSLAGVAVMLVMPVN 480
QY 481 AVAMAKTKTYQVAHMKSKDNRIKLWNEILNGIKVLKLVAVELAFKDKVLAIROBELKVLK 540
DB 481 AVAMAKTKTYQVAHMKSKDNRIKLWNEILNGIKVLKLVAVELAFKDKVLAIROBELKVLK 540
QY 541 KSAYLSAVGTFTVCTPPELVALCTFAVYVITDENNILDQAQTAFAVSLALFNILRFPNLILP 600
DB 541 KSAYLSAVGTFTVCTPPELVALCTFAVYVITDENNILDQAQTAFAVSLALFNILRFPNLILP 600
QY 601 WYISSIVQASVLSKRLRFLPSHEELEPDSIERPPKDGCGTNSITVRNATFTWASDPT 660
DB 601 WYISSIVQASVLSKRLRFLPSHEELEPDSIERPPKDGCGTNSITVRNATFTWASDPT 660
QY 661 LNGITFTSIPGALVAVGVGCGKSSLSALLAEKDVGHVAIKGSVAYVPOQAWIQND 720
DB 661 LNGITFTSIPGALVAVGVGCGKSSLSALLAEKDVGHVAIKGSVAYVPOQAWIQND 720
QY 721 SLRENILFGCQLEPYRSVIOACALLPDLPLPSGDRTEIGEKGWNLGGQKQVSLAR 780

DB 721 SLRENILFGCQLEPYRSVIOACALLPDLPLPSGDRTEIGEKGWNLGGQKQVSLAR 780
QY 781 AVYSNADIYLPDDPLSAVDHVGKHI FENVIGPKGMLKNKTRILVTHSMYLPQVDVLIIV 840
DB 781 AVYSNADIYLPDDPLSAVDHVGKHI FENVIGPKGMLKNKTRILVTHSMYLPQVDVLIIV 840
QY 841 MSGGKISEMGSYQELLARDGAFAEFLRTYASTEQQDAEENGVTGVSFGPKGKAKQWNGM 900
DB 841 MSGGKISEMGSYQELLARDGAFAEFLRTYASTEQQDAEENGVTGVSFGPKGKAKQWNGM 900
QY 901 LVTSAGLQORQLSSSSSYSGDISRHNSTAELOKABAKKEETWKLMEADKAQTQGVKL 960
DB 901 LVTSAGLQORQLSSSSSYSGDISRHNSTAELOKABAKKEETWKLMEADKAQTQGVKL 960
QY 961 SVYDYMKAIGLIFISFLSIFLPMCNHVSALASNYWLSLWTDPI VNGTQEHKTVRLSVYG 1020
DB 961 SVYDYMKAIGLIFISFLSIFLPMCNHVSALASNYWLSLWTDPI VNGTQEHKTVRLSVYG 1020
QY 1021 ALGISQGIAGFYGSMAVISGIGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFSKEL 1080
DB 1021 ALGISQGIAGFYGSMAVISGIGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFSKEL 1080
QY 1081 DTVDSMIDPEVIMKPMGSLFNIVIGACIVILLATPIAAIIIPBLGLIYFFVQRFYVASSRQL 1140
DB 1081 DTVDSMIDPEVIMKPMGSLFNIVIGACIVILLATPIAAIIIPBLGLIYFFVQRFYVASSRQL 1140
QY 1141 KRLESVSRSPPVYSHNETLLGVSVIRAFEEQERFIHQSDLKVDENQKAYYPSIIVANRWLA 1200
DB 1141 KRLESVSRSPPVYSHNETLLGVSVIRAFEEQERFIHQSDLKVDENQKAYYPSIIVANRWLA 1200
QY 1201 VRLECVGNCIVLFAALFAVI SRHSLSAGLVGLSVSYSLQVTTYLNMLVRMSSEMETNIVA 1260
DB 1201 VRLECVGNCIVLFAALFAVI SRHSLSAGLVGLSVSYSLQVTTYLNMLVRMSSEMETNIVA 1260
QY 1261 VERLKEYSETKEAPWQIQETAPPSWQVGRVEFRNTCLARYEDLDLDFVLRHINVTINGG 1320
DB 1261 VERLKEYSETKEAPWQIQETAPPSWQVGRVEFRNTCLARYEDLDLDFVLRHINVTINGG 1320
QY 1321 EKVGI VGRGTAGKSSLTGLFRINESARGEI IIDGINIAKIGLHDLRFKTIIPDDPVLF 1380
DB 1321 EKVGI VGRGTAGKSSLTGLFRINESARGEI IIDGINIAKIGLHDLRFKTIIPDDPVLF 1380
QY 1381 SGLSRMNLDPFSQYSDEEVWTSLELAHLKDFVSALPDKLDHECAGGENLSVGQRLVCL 1440
DB 1381 SGLSRMNLDPFSQYSDEEVWTSLELAHLKDFVSALPDKLDHECAGGENLSVGQRLVCL 1440
QY 1441 ARALLRKTIIIVLDEATAVDETDLLIQSTIRQTFEDCTVLTIAHRLNTIMDYTRVIVL 1500
DB 1441 ARALLRKTIIIVLDEATAVDETDLLIQSTIRQTFEDCTVLTIAHRLNTIMDYTRVIVL 1500
QY 1501 DKGETOEYGA PSDLLQORGLFVSMADAGLV 1531
DB 1501 DKGETOEYGA PSDLLQORGLFVSMADAGLV 1531

RESULT 3
US-08-460-907B-4
; Sequence 4, Application US/08460907B
; Patent No. 5891724
; GENERAL INFORMATION:
; APPLICANT: Deeley, Roger G.
; APPLICANT: Cole, Susan P.C.
; TITLE OF INVENTION: METHODS FOR CONFERRING MULTIDRUG
; RESISTANCE ON A CELL
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PAREQ RESEARCH & DEVELOPMENT INNOVATIONS
; STREET: Queen's University at Kingston
; CITY: Kingston
; STATE: Ontario
; COUNTRY: CANADA
; ZIP: K7L 3N6

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; FILING DATE: 05-JUN-1995
; APPLICATION NUMBER: US/08/460,907B
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; FILING DATE: 07/966,923
; FILING DATE: 27-OCT-1992
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/029,340
; FILING DATE: 8-MAR-1993
; CLASSIFICATION: 424
; APPLICATION DATA:
; FILING DATE: 08/141,893
; FILING DATE: 26-OCT-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/407,207
; FILING DATE: 20-MAR-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Steeg, Carol Miernicki
; REGISTRATION NUMBER: 39,539
; REFERENCE/DOCKET NUMBER: Q1551
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (613) 545-2342
; TELEFAX: (613) 545-6853
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1531 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-460-907B-4

Query Match 79.4%; Score 7860; DB 2; Length 1531;
Besc Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCADGSDPLMDNNTWNTSNPDFTKCFQNTVLVWVPCFYLWACFPFFLYLSRH 60
DB 1 MALRGFCADGSDPLMDNNTWNTSNPDFTKCFQNTVLVWVPCFYLWACFPFFLYLSRH 60
QY 61 DRGIQMTPLNKTALGFLWIVCWADLFYSFWERSRGI FLAPVFLVSPPTLLGTTLLA 120
DB 61 DRGIQMTPLNKTALGFLWIVCWADLFYSFWERSRGI FLAPVFLVSPPTLLGTTLLA 120
QY 121 TFLQLERKGVQSSGIMLTFMLVALVCALAILRSKIMTALKEDAQVDLPDITFYVYFS 180
DB 121 TFLQLERKGVQSSGIMLTFMLVALVCALAILRSKIMTALKEDAQVDLPDITFYVYFS 180
QY 181 LLLQLVLSCFSDRSLPSETIHPNCPCESSASFSLRITFWITGLIVRGYRPLEGSD 240
DB 181 LLLQLVLSCFSDRSLPSETIHPNCPCESSASFSLRITFWITGLIVRGYRPLEGSD 240
QY 241 LWSLNKETSQVVPVLVKNWKECAKTRKQPVKVYSSKDPAPQKSSKVDANEVEAL 300
DB 241 LWSLNKETSQVVPVLVKNWKECAKTRKQPVKVYSSKDPAPQKSSKVDANEVEAL 300
QY 301 IVKSPQKEWNSLFLKVLKTFPGPYFLMSFFPKAIDHLMWFSGPOILKLLIKFVNDTKAPD 360
DB 301 IVKSPQKEWNSLFLKVLKTFPGPYFLMSFFPKAIDHLMWFSGPOILKLLIKFVNDTKAPD 360
QY 361 WQGYFYTVLLFVTAQLOTLVLHQYFHCIFVSGMRKTAIVIGAVYRKALVITNSARKSSTV 420
DB 361 WQGYFYTVLLFVTAQLOTLVLHQYFHCIFVSGMRKTAIVIGAVYRKALVITNSARKSSTV 420
QY 421 GEIVNLSVDAQRFDMLATYINMIWSAPLQVILALYLLNLGSPSVLAGVAVMLVMPVN 480

```

Db 1501 DKGEIQEYCAPSDLLQOORGLFYSMKADAGLV 1531

RESULT 4  
US-08-463-179A-4  
; Sequence 4, Application US/08463179A  
; Patent No. 6001563  
; GENERAL INFORMATION:  
; APPLICANT: Cole, Susan P.C.  
; APPLICANT: Deeley, Roger G.  
; TITLE OF INVENTION: METHODS FOR IDENTIFYING CHEMOSENSITIZERS  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street, suite 510  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/463,179A  
; FILING DATE:  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/966,923  
; FILING DATE: 27-OCT-1992  
; APPLICATION NUMBER: 08/029,340  
; FILING DATE: 8-MAR-1993  
; APPLICATION NUMBER: 08/141,893  
; FILING DATE: 26-OCT-1993  
; APPLICATION NUMBER: 08/407,207  
; FILING DATE: 20-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: DeConti, Giulio A. Jr.  
; REGISTRATION NUMBER: 31,503  
; REFERENCE/DOCKET NUMBER: PQI-002CP8  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 227-7400  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1531 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-463-179A-4

Query Match 79.4%; Score 7860; DB 3; Length 1531;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCSADGSDPLMDNVTWNTSNPDFTKCFQNTLVWVPCFYLMACFPFYLYLSRH 60  
DB 1 MALRGFCSADGSDPLMDNVTWNTSNPDFTKCFQNTLVWVPCFYLMACFPFYLYLSRH 60  
QY 61 DRGYIQMTPLNKTALGFLLMIVCWADLFYFWSRSGIFLAPVFLVSPDLLGTTLLA 120  
DB 61 DRGYIQMTPLNKTALGFLLMIVCWADLFYFWSRSGIFLAPVFLVSPDLLGTTLLA 120  
QY 121 TFLIQLERRKGVOSSGIMLTFWLVALCALALILRSKIMTALKEDAQVDLFRDITFYVYFS 180  
DB 121 TFLIQLERRKGVOSSGIMLTFWLVALCALALILRSKIMTALKEDAQVDLFRDITFYVYFS 180  
QY 181 LLLIQLVLSFCSDRSPFLSETIHDNPPCPSSASFLSRITTFWITGLIVRGYRQPLEGSD 240  
DB 181 LLLIQLVLSFCSDRSPFLSETIHDNPPCPSSASFLSRITTFWITGLIVRGYRQPLEGSD 240  
QY 241 LWSLNKEDTSEQVVPVLVKNMKECAKTRKQPVKVYSSKOPAKPSSKVDANEVEAL 300

Db 241 LWSLNKEDTSEQVVPVLVKNMKECAKTRKQPVKVYSSKOPAKPSSKVDANEVEAL 300  
QY 301 IVKSPQKWNPSLFLKVLKTYKTFPGPYFLMSFFPKAIHDLMMFSGPOLIKLLIKFVNDTKAPD 360  
Db 301 IVKSPQKWNPSLFLKVLKTYKTFPGPYFLMSFFPKAIHDLMMFSGPOLIKLLIKFVNDTKAPD 360  
QY 361 WQGYFYTVLLFVTACLOTLVLHQYPHICFVSGMRIKTAVIGAVYRKALVITNSARKSSTV 420  
Db 361 WQGYFYTVLLFVTACLOTLVLHQYPHICFVSGMRIKTAVIGAVYRKALVITNSARKSSTV 420  
QY 421 GEIVNLSMSVDAQRFMDLATYINMIWSAPLOVILALYLILNLGSPSVLAGVAVMLMVPYN 480  
Db 421 GEIVNLSMSVDAQRFMDLATYINMIWSAPLOVILALYLILNLGSPSVLAGVAVMLMVPYN 480  
QY 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKLYAWELAFKDKVLAIQEBELKVLK 540  
Db 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKLYAWELAFKDKVLAIQEBELKVLK 540  
QY 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTTIDENNILDAQTAFAVSLALFNILRFPNLILP 600  
Db 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTTIDENNILDAQTAFAVSLALFNILRFPNLILP 600  
QY 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDCGGTNSITVRNATFTWARSDDPT 660  
Db 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDCGGTNSITVRNATFTWARSDDPT 660  
QY 661 LMGITFSIPEGALVAVVGQVCGKSSLLSALLAEMDKVEGHVAIKGSVAYVPPQQAOWND 720  
Db 661 LMGITFSIPEGALVAVVGQVCGKSSLLSALLAEMDKVEGHVAIKGSVAYVPPQQAOWND 720  
QY 721 SLRENILFGCOLEEBPYRSVIQACALLPDLLEIPSGDRTEIGEKGVNLSGGQKQVSLAR 780  
Db 721 SLRENILFGCOLEEBPYRSVIQACALLPDLLEIPSGDRTEIGEKGVNLSGGQKQVSLAR 780  
QY 781 AVYSNADIYLFDDPLSAVDAHVGHKIPENVIGPKGMLKNKTRILVTHSMYSYLPQVDVIIV 840  
Db 781 AVYSNADIYLFDDPLSAVDAHVGHKIPENVIGPKGMLKNKTRILVTHSMYSYLPQVDVIIV 840  
QY 841 MSGGKI SEMGSYQELLARDGAFABFLRTVASTEQDABENGVTGSGPGKEAKOMENGM 900  
Db 841 MSGGKI SEMGSYQELLARDGAFABFLRTVASTEQDABENGVTGSGPGKEAKOMENGM 900  
QY 901 LVTDSAGKQLOQLSSSSSYSGDISRHHNSTAELOKAEAKKEETWKLMEADKAQTCQVKL 960  
Db 901 LVTDSAGKQLOQLSSSSSYSGDISRHHNSTAELOKAEAKKEETWKLMEADKAQTCQVKL 960  
QY 961 SVYWDYMKAIGLFISFLSIFLPMCHVUSALASNYWLSLWTDPIVNGTQEHTKVRLSVYG 1020  
Db 961 SVYWDYMKAIGLFISFLSIFLPMCHVUSALASNYWLSLWTDPIVNGTQEHTKVRLSVYG 1020  
QY 1021 ALGISQGI AVFGYSMAVSI GIGILASRCLVOLLHLSILRSPMSFFERTPSGNLVNRFPSKEL 1080  
Db 1021 ALGISQGI AVFGYSMAVSI GIGILASRCLVOLLHLSILRSPMSFFERTPSGNLVNRFPSKEL 1080  
QY 1081 DTVDSMIPEVIKMFMSGLFNIVIGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQL 1140  
Db 1081 DTVDSMIPEVIKMFMSGLFNIVIGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQL 1140  
QY 1141 KRLSVSRSPYVSHNETLLGVSVIRAPEOERFIHQSDLKVDENQKAYPSIVANRWLA 1200  
Db 1141 KRLSVSRSPYVSHNETLLGVSVIRAPEOERFIHQSDLKVDENQKAYPSIVANRWLA 1200  
QY 1201 VRLCEVGNICVILFAALFAVISHRSISAGLVLSYSISQVTTYLNLWLRMSSEMETNIVA 1260  
Db 1201 VRLCEVGNICVILFAALFAVISHRSISAGLVLSYSISQVTTYLNLWLRMSSEMETNIVA 1260  
QY 1261 VERLKEYSETEKEAPWQIQETAPSSWPQVGRVFRNYCLARYREDLDFVLRHINTYINGG 1320  
Db 1261 VERLKEYSETEKEAPWQIQETAPSSWPQVGRVFRNYCLARYREDLDFVLRHINTYINGG 1320  
QY 1321 EKVGI VGTGAGKSLTLGLFRINESAGEIIIDGINIAKIGLHDLRPFKITIIPDPVLPF 1380

Db 1321 EKVGIVGRTGAGKSSLTGLFRINESAGEIITDGINIAKIGHDLRFKTIIPQDPVLF 1380  
 Qy 1381 SGLRNWLPFSQYSDDEEYVWTSLELAHLKDFVSALPDKLDHCAEGENLSVGQRLVCL 1440  
 Db 1381 SGLRNWLPFSQYSDDEEYVWTSLELAHLKDFVSALPDKLDHCAEGENLSVGQRLVCL 1440  
 Qy 1441 ARALLKTKILVLDATAVDDLEDDLIQSTIRTOPEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 Db 1441 ARALLKTKILVLDATAVDDLEDDLIQSTIRTOPEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 Qy 1501 DKGEIQEYCAPSDLLQORGLFYSMAKDAGLV 1531  
 Db 1501 DKGEIQEYCAPSDLLQORGLFYSMAKDAGLV 1531

RESULT 5  
 US-08-461-384B-4  
 ; Sequence 4, Application US/08461384B  
 ; Patent No. 6025473  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cole, Susan P.C.  
 ; APPLICANT: Deesley, Roger G.  
 ; TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS  
 ; NUMBER OF SEQUENCES: 10  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS  
 ; STREET: Queen's University at Kingston  
 ; CITY: Kingston  
 ; STATE: Ontario  
 ; COUNTRY: CANADA  
 ; ZIP: K7L 3N6  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: ASCII text  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/461,384B  
 ; FILING DATE: 05-JUN-95  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 07/966,923  
 ; FILING DATE: 27-OCT-1992  
 ; APPLICATION NUMBER: 08/029,340  
 ; FILING DATE: 8-MAR-1993  
 ; APPLICATION NUMBER: 08/141,893  
 ; FILING DATE: 26-OCT-1993  
 ; APPLICATION NUMBER: 08/407,207  
 ; FILING DATE: 20-MAR-1995  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Steeg, Carol Miernicki  
 ; REGISTRATION NUMBER: 39,539  
 ; REFERENCE/DOCKET NUMBER: Q1547  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (613) 545-2342  
 ; TELEFAX: (613) 545-6853  
 ; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1531 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 US-08-461-384B-4

Query Match 79.4%; Score 7860; DB 3; Length 1531;  
 Best Local Similarity 100.08; Pred No. 0;  
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 MALRGFCADGSDPLWDDNVTWNTSNPDFTKCFQNTLVWVPCFYLWACFPFFFLYLSRH 60  
 Db 1 MALRGFCADGSDPLWDDNVTWNTSNPDFTKCFQNTLVWVPCFYLWACFPFFFLYLSRH 60  
 Qy 61 DRGVIQMTPLNKTALGFLWIVCWADLFYSFWERSRGIFLAPVFLVSPILLGITTLLA 120

Db 61 DRGVIQMTPLNKTALGFLWIVCWADLFYSFWERSRGIFLAPVFLVSPILLGITTLLA 120  
 Qy 121 TFLIQLERRKGVOSSGIMLTFWLVALVCALAILRSKIMTALKEDAQVDFRDIITYVYVFS 180  
 Db 121 TFLIQLERRKGVOSSGIMLTFWLVALVCALAILRSKIMTALKEDAQVDFRDIITYVYVFS 180  
 Qy 181 LLLIQVLVSCFSDRSPFLSETHDPNCPSESSASPLSRITFWWITGLIVRGVROPLEGSD 240  
 Db 181 LLLIQVLVSCFSDRSPFLSETHDPNCPSESSASPLSRITFWWITGLIVRGVROPLEGSD 240  
 Qy 241 LWSLNKEDTSEQVVPVLVKNWKKCECAKTRKQPVKVYVYSSKDPAPQKSSKVDANEVEAL 300  
 Db 241 LWSLNKEDTSEQVVPVLVKNWKKCECAKTRKQPVKVYVYSSKDPAPQKSSKVDANEVEAL 300  
 Qy 301 IVKSPQKEMNPSLFKVLTKTGPFLMSFFPKA IHDLMFSGPQILKLLIKFVNDTKAPD 360  
 Db 301 IVKSPQKEMNPSLFKVLTKTGPFLMSFFPKA IHDLMFSGPQILKLLIKFVNDTKAPD 360  
 Qy 361 WQGYFYTVLLFWTACLTQLVHLQYFHCFSVSGMRIKTAIVIGAVYRKALVITNSARKSSTV 420  
 Db 361 WQGYFYTVLLFWTACLTQLVHLQYFHCFSVSGMRIKTAIVIGAVYRKALVITNSARKSSTV 420  
 Qy 421 GEIVNLSVDAQREFMDLATYINMIWSAPLOQVILALYLLMLNLGSPSVLAGVAVMLVMPVN 480  
 Db 421 GEIVNLSVDAQREFMDLATYINMIWSAPLOQVILALYLLMLNLGSPSVLAGVAVMLVMPVN 480  
 Qy 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAWELAPKDKVLAIROBELKVLYK 540  
 Db 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAWELAPKDKVLAIROBELKVLYK 540  
 Qy 541 KSAVLSAVGTFTWCTPFLVALCTFAVVVTIDENNILDAQTAFAVLSLAFNLIRPLNLTP 600  
 Db 541 KSAVLSAVGTFTWCTPFLVALCTFAVVVTIDENNILDAQTAFAVLSLAFNLIRPLNLTP 600  
 Qy 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGGNSITVRNATFTWARSPPPT 660  
 Db 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGGNSITVRNATFTWARSPPPT 660  
 Qy 661 LNGITFSTPEGALVAVVGQCGKSSLSALLAEMDKVEGHVAKGSVAVYVPOQAWTQND 720  
 Db 661 LNGITFSTPEGALVAVVGQCGKSSLSALLAEMDKVEGHVAKGSVAVYVPOQAWTQND 720  
 Qy 721 SLRENILFGQLEERPYRSVIQACALLPDLLEILPSGDRTEIGEKGVLGSGQKQKRVSLAR 780  
 Db 721 SLRENILFGQLEERPYRSVIQACALLPDLLEILPSGDRTEIGEKGVLGSGQKQKRVSLAR 780  
 Qy 781 AVYSNADIYLPDDPLSAVDAAHVGHKIPENVIQPKGMLKNKTRILVTHSMSYLPQVDVIV 840  
 Db 781 AVYSNADIYLPDDPLSAVDAAHVGHKIPENVIQPKGMLKNKTRILVTHSMSYLPQVDVIV 840  
 Qy 841 MSGGKISEMGSYQELLARDGAFAEFLRTYASTEQDAEENGVTGVSQPGKEAKOMENGM 900  
 Db 841 MSGGKISEMGSYQELLARDGAFAEFLRTYASTEQDAEENGVTGVSQPGKEAKOMENGM 900  
 Qy 901 LVTDSAGKQLQRLSSSSSYSGDISRHHNSTAEIQLKAEAKKEETWKLMEADKATQGVKL 960  
 Db 901 LVTDSAGKQLQRLSSSSSYSGDISRHHNSTAEIQLKAEAKKEETWKLMEADKATQGVKL 960  
 Qy 961 SVYWDYMKAIGLFISFLSIFLFCMNHVSALASYNWLSLTDDDPVINGTQHTKVLRSVYG 1020  
 Db 961 SVYWDYMKAIGLFISFLSIFLFCMNHVSALASYNWLSLTDDDPVINGTQHTKVLRSVYG 1020  
 Qy 1021 ALGISQGIAGVFGSMAYSIGIILASRCLHVDLLHSILRSWSPFPPTSGNLVNRFSKEL 1080  
 Db 1021 ALGISQGIAGVFGSMAYSIGIILASRCLHVDLLHSILRSWSPFPPTSGNLVNRFSKEL 1080  
 Qy 1081 DTVDSMIPEVIMKPMGSLFNIVIGACIVILLATPIAAIIIPPLGLIYFFQRFYVASSRQL 1140  
 Db 1081 DTVDSMIPEVIMKPMGSLFNIVIGACIVILLATPIAAIIIPPLGLIYFFQRFYVASSRQL 1140  
 Qy 1141 KRLESVSRSPVYSHFNETHLGVSVIRAFERQERFIHQSDLKVDENQKAYVPSIVANRWLA 1200  
 Db 1141 KRLESVSRSPVYSHFNETHLGVSVIRAFERQERFIHQSDLKVDENQKAYVPSIVANRWLA 1200

[illegible]

```

RESULT 6
US-09-647-140B-19
; Sequence 19, Application US/09647140B
; Patent No. 6803184
; GENERAL INFORMATION:
; APPLICANT: Fox Chase Cancer Center
; APPLICANT: Kruh, Gary D.
; APPLICANT: Lee, Kun
; APPLICANT: Belinsky, Martin G.
; APPLICANT: Bain, Lisa J.
; TITLE OF INVENTION: MRP-Related ABC Transporter Encoding
; TITLE OF INVENTION: Nucleic Acids and Methods of Use Thereof
; FILE REFERENCE: FCCC 98-02
; CURRENT APPLICATION NUMBER: US/09/647,140B
; CURRENT FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: PCT/US99/06644
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 60/079,759
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/095,153
; PRIOR FILING DATE: 1998-08-03
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 19
; LENGTH: 1531
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-647-140B-19

```

[illegible]

181	Db	LLLLQLVLSCFSDRSPLFSETHDNPCCPSSSASFLSRITFFWITGLI VRGVTRQPLEGSD	240
241	Qy	LWSLNKEDTSEQVVPVLVRXWKKCECAKTRKQPVKVYVYSSKDPAQPKSSKVDANEVEAL	300
241	Db	LWSLNKEDTSEQVVPVLVRXWKKCECAKTRKQPVKVYVYSSKDPAQPKSSKVDANEVEAL	300
301	Qy	IVKSPKQEWNPSLFKVLKYKTFGYPFLMSFFPKAIHDLMMFSGPOILKLLIKFVNDTKAPD	360
301	Db	IVKSPKQEWNPSLFKVLKYKTFGYPFLMSFFPKAIHDLMMFSGPOILKLLIKFVNDTKAPD	360
361	Qy	WQGYFYTVLLFTVTAQLQTLVLHGYPHICFVSGWRIKTAIVGAYVRKALVTITNSARKSSTV	420
361	Db	WQGYFYTVLLFTVTAQLQTLVLHGYPHICFVSGWRIKTAIVGAYVRKALVTITNSARKSSTV	420
421	Qy	GEIVNLSMSVDAQRFMDLATYINNIWSAPLQVILALYLLMLNLGSPSVLAGVAVMVLMPVN	480
421	Db	GEIVNLSMSVDAQRFMDLATYINNIWSAPLQVILALYLLMLNLGSPSVLAGVAVMVLMPVN	480
481	Qy	AVNAMTKTYQVAHMKSKONRIKLMNEILNGIKVLKLYAWELAFKDKVLAIROEELKVLK	540
481	Db	AVNAMTKTYQVAHMKSKONRIKLMNEILNGIKVLKLYAWELAFKDKVLAIROEELKVLK	540
541	Qy	KSAYLSAVGTFTWCTPEIVALCTPAVYVTIDENNILDAQTA FVSIALFNILRFPNLNLP	600
541	Db	KSAYLSAVGTFTWCTPEIVALCTPAVYVTIDENNILDAQTA FVSIALFNILRFPNLNLP	600
601	Qy	MVSISSVQASVSLKRLRIIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWARS DPPT	660
601	Db	MVSISSVQASVSLKRLRIIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWARS DPPT	660
661	Qy	LNGITPSIPEGALVAVVGQVCGCKSILLSALLAEMDKVEGHVAIKGSVAYVPOQAWIQND	720
661	Db	LNGITPSIPEGALVAVVGQVCGCKSILLSALLAEMDKVEGHVAIKGSVAYVPOQAWIQND	720
721	Qy	SLRENILFGCOLPEEPYRSVIOACALLPDLIELPSGDRTEIGBKGNLSGGQKQRVSLAR	780
721	Db	SLRENILFGCOLPEEPYRSVIOACALLPDLIELPSGDRTEIGBKGNLSGGQKQRVSLAR	780
781	Qy	AVYSNADIYFDDPLSAVDHVGKHFENVIGPKMLKNKTRILVTHSHSYLSPQVDVIIV	840
781	Db	AVYSNADIYFDDPLSAVDHVGKHFENVIGPKMLKNKTRILVTHSHSYLSPQVDVIIV	840
841	Qy	MSGGKISEMGSYQELLARDGAFABFURTYASTEQEQDQDAEENGVTGVS GPGKEAKQWENG M	900
841	Db	MSGGKISEMGSYQELLARDGAFABFURTYASTEQEQDQDAEENGVTGVS GPGKEAKQWENG M	900
901	Qy	LVWDSAGKQLOROLSSSSVSGDISRHNSSTAELQKAEAKKEETWKLMEADKAQTGOVKL	960
901	Db	LVWDSAGKQLOROLSSSSVSGDISRHNSSTAELQKAEAKKEETWKLMEADKAQTGOVKL	960
961	Qy	SVTWDMYKAIGLPIFLSIFLPMCNHVSALASNYLSLWTD DDPINVNGTOBHTKVRLSVYG	1020
961	Db	SVTWDMYKAIGLPIFLSIFLPMCNHVSALASNYLSLWTD DDPINVNGTOBHTKVRLSVYG	1020
1021	Qy	ALGISOGIAVFGYSMAVSI GGILLASRCLHVDDLHSILRSPMSFFERTP SGNLVNRFSKEL	1080
1021	Db	ALGISOGIAVFGYSMAVSI GGILLASRCLHVDDLHSILRSPMSFFERTP SGNLVNRFSKEL	1080
1081	Qy	DTVDSMIPVIMKFMGSLFNVIACACIVILLATPIAAIIIPPLGLIYFFVORFVYASSROL	1140
1081	Db	DTVDSMIPVIMKFMGSLFNVIACACIVILLATPIAAIIIPPLGLIYFFVORFVYASSROL	1140
1141	Qy	KRLIESRSRSPVYSHNETLLGVSVIRAFBEOERFIHQSDLKVDENOKAYYPSIVANRWLA	1200
1141	Db	KRLIESRSRSPVYSHNETLLGVSVIRAFBEOERFIHQSDLKVDENOKAYYPSIVANRWLA	1200
1201	Qy	VRLECVGNCIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTYTLNMLVRMSSEMETNIVA	1260
1201	Db	VRLECVGNCIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTYTLNMLVRMSSEMETNIVA	1260
1261	Qy	VERLKEYSETEKEAPQWIOETAPPSWPQVGRVEFRNYCLURYEDDLDFVLRHNVINGG	1320
1261	Db	VERLKEYSETEKEAPQWIOETAPPSWPQVGRVEFRNYCLURYEDDLDFVLRHNVINGG	1320

QY 1321 EKVGIVGRTGAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKXITIIPODPVLF 1380  
 DB 1321 EKVGIVGRTGAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKXITIIPODPVLF 1380  
 QY 1381 SGLSRNLDPFQSDSEEVWTSLELAHLKDFVSALPDKLDHECAEGENLSVGQRLVCL 1440  
 DB 1381 SGLSRNLDPFQSDSEEVWTSLELAHLKDFVSALPDKLDHECAEGENLSVGQRLVCL 1440  
 QY 1441 ABALLRKTILVLDEATAVDLETDLLIOSTTRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 DB 1441 ABALLRKTILVLDEATAVDLETDLLIOSTTRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 QY 1501 DKGEIQEYCAPSDLLQOORGLFYSMADAGLV 1531  
 DB 1501 DKGEIQEYCAPSDLLQOORGLFYSMADAGLV 1531

## RESULT 7

US-08-141-893-2  
 ; Sequence 2, Application US/08141893  
 ; Patent No. 5489519  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cole, Susan P.C.  
 ; APPLICANT: Deeley, Roger G.  
 ; TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEIN  
 ; NUMBER OF SEQUENCES: 2  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSES: LAHIVE & COCKFIELD  
 ; STREET: 60 STATE STREET, SUITE 510  
 ; CITY: BOSTON  
 ; STATE: MASSACHUSETTS  
 ; COUNTRY: USA  
 ; ZIP: 02109  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/141,893  
 ; FILING DATE:  
 ; CLASSIFICATION: 424  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 07/966,923; 08/029,340  
 ; FILING DATE: 27-OCT-1992; 8-MAR-1993  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Deconti, Giulio A. Jr.  
 ; REGISTRATION NUMBER: 31,503  
 ; REFERENCE/DOCKET NUMBER: PQI-002  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (617) 227-7400  
 ; TELEFAX: (617) 227-5149  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1531 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 US-08-141-893-2

Query Match 79.3%; Score 7849; DB 1; Length 1531;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 MALRGCSDGSDPLDWDNVNTNSNPDFTKCFQNTVLVWVPCFVLWACFPFFLYLSRH 60  
 DB 1 MALRGCSDGSDPLDWDNVNTNSNPDFTKCFQNTVLVWVPCFVLWACFPFFLYLSRH 60  
 QY 61 DRGIQMTPLNKTALGFLWVWCVADLFYSFWRSGIFLAPVFLVSPFLIGITLLA 120  
 DB 61 DRGIQMTPLNKTALGFLWVWCVADLFYSFWRSGIFLAPVFLVSPFLIGITLLA 120

QY 121 TFLIOLERRKGVOSSGIMLTFWLVALVCALAILRSKINTALKEDAQVDLFRDITPVYVFS 180  
 DB 121 TFLIOLERRKGVOSSGIMLTFWLVALVCALAILRSKINTALKEDAQVDLFRDITPVYVFS 180  
 QY 181 LLLIQLVLSGSDRSPLESETHDNPCESSASFLSRITFWWITGLIVRGVQPLEGSD 240  
 DB 181 LLLIQLVLSGSDRSPLESETHDNPCESSASFLSRITFWWITGLIVRGVQPLEGSD 240  
 QY 241 LWSLNKEDTSQVVPVLVKNWKCACAKTRKQPVKVYSSKDPAPQPKESKVDANEVEAL 300  
 DB 241 LWSLNKEDTSQVVPVLVKNWKCACAKTRKQPVKVYSSKDPAPQPKESKVDANEVEAL 300  
 QY 301 IVKSPQKEMNLSLKVLYKTFPGYFLMSFFPKAIDHLMWFGSPQILKLLIKFVNDTKAPD 360  
 DB 301 IVKSPQKEMNLSLKVLYKTFPGYFLMSFFPKAIDHLMWFGSPQILKLLIKFVNDTKAPD 360  
 QY 361 WQGYFYTVLLFVTTACLOTVLHQLVHFICFVSGMRIKTAVIGAVYRKALVITNSARKSSTV 420  
 DB 361 WQGYFYTVLLFVTTACLOTVLHQLVHFICFVSGMRIKTAVIGAVYRKALVITNSARKSSTV 420  
 QY 421 GEIVNLMSVDAQRFMDLATYINMIWSAPLQVILALYLLWNLGSPSVLAGVAVMLMVPVN 480  
 DB 421 GEIVNLMSVDAQRFMDLATYINMIWSAPLQVILALYLLWNLGSPSVLAGVAVMLMVPVN 480  
 QY 481 AVMAKTKTYOVAMHMKONRIKLMNETLNGIKVLKLYAWELAFKDKVLATROBELKVLK 540  
 DB 481 AVMAKTKTYOVAMHMKONRIKLMNETLNGIKVLKLYAWELAFKDKVLATROBELKVLK 540  
 QY 541 KSAYLSAVGTFTWCTPFLVALCTFAVVTIDENNILDAQTAFAVSLALFNILRPLNLPL 600  
 DB 541 KSAYLSAVGTFTWCTPFLVALCTFAVVTIDENNILDAQTAFAVSLALFNILRPLNLPL 600  
 QY 601 MVISSIQAQSVSLKRLRIFLSHELEPDSIPRRPVKCGGTSITVRNATTWARSDDPT 660  
 DB 601 MVISSIQAQSVSLKRLRIFLSHELEPDSIPRRPVKCGGTSITVRNATTWARSDDPT 660  
 QY 661 LNGITFSIPEGALVAVVGVCGCKSLLSALLAEMDKVEGHVAKGVSVAVVPQQAOWND 720  
 DB 661 LNGITFSIPEGALVAVVGVCGCKSLLSALLAEMDKVEGHVAKGVSVAVVPQQAOWND 720  
 QY 721 SLRENILFGCOLEBPYRSVIQACALLPDLLEILPSGDRTEIGEGVNLSSGQKQVSLAR 780  
 DB 721 SLRENILFGCOLEBPYRSVIQACALLPDLLEILPSGDRTEIGEGVNLSSGQKQVSLAR 780  
 QY 781 AVYSNADIYLFDDPLSAVDAHVGHIENVIKPGMLKNKTRILVTHSMSVLPQVDVLIIV 840  
 DB 781 AVYSNADIYLFDDPLSAVDAHVGHIENVIKPGMLKNKTRILVTHSMSVLPQVDVLIIV 840  
 QY 841 MSGGKISEMGSYQELLARDGAFAPAEFLTYASTEQDAEENGVTGVSQPGKEAKQEMNGM 900  
 DB 841 MSGGKISEMGSYQELLARDGAFAPAEFLTYASTEQDAEENGVTGVSQPGKEAKQEMNGM 900  
 QY 901 LVTDGAKQLQRLSSSSSYSGDISRHHNSTAEILOKAEKBEETWKLMEADKAQTOGVKL 960  
 DB 901 LVTDGAKQLQRLSSSSSYSGDISRHHNSTAEILOKAEKBEETWKLMEADKAQTOGVKL 960  
 QY 961 SVYWDYMKAGLFIISFLSIIFLCNHNVSALASVNLSTDDPIVNGTOHTKVLRSVYG 1020  
 DB 961 SVYWDYMKAGLFIISFLSIIFLCNHNVSALASVNLSTDDPIVNGTOHTKVLRSVYG 1020  
 QY 1021 ALGISQGIAGVFGYSMAVSIIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFPSKEL 1080  
 DB 1021 ALGISQGIAGVFGYSMAVSIIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFPSKEL 1080  
 QY 1081 DTVDSDMIPVIMKPMGSLFNVIAGACIVILLATPIAAIIIPPLGLIYFFVORFYVASSRQL 1140  
 DB 1081 DTVDSDMIPVIMKPMGSLFNVIAGACIVILLATPIAAIIIPPLGLIYFFVORFYVASSRQL 1140  
 QY 1141 KRLESVSRSPVYSHFNETLGLVSVIRAFERQERFIHOSDLKVDENQKAYYPSIVANRWLA 1200  
 DB 1141 KRLESVSRSPVYSHFNETLGLVSVIRAFERQERFIHOSDLKVDENQKAYYPSIVANRWLA 1200  
 QY 1201 VRLECVGNCIVLFAALFAVISRHSLSAGLVLSVSYSLQVTTYLNLWLVNRMSSMETNIVA 1260

```

Db 1201 VLECVGNCIVLFAALFAVISHSLAGLVSVSYLQVTTYNLWVMSSEMTNIVA 1260
QY 1261 VERLKEYSETEKEAPWQIQTETAPPSSWQVGRVFRNRYCLAYREDLDVLRHINTINGG 1320
Db 1261 VERLKEYSETEKEAPWQIQTETAPPSSWQVGRVFRNRYCLAYREDLDVLRHINTINGG 1320
QY 1321 EKVGVTRGTGAGKSLTLGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
Db 1321 EKVGVTRGTGAGKSLTLGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
QY 1381 SGLSRMLNDPPFQSYDEEVTWTSLELAHLKDFVSALPKDLHDECARGENLSVGQRLVCL 1440
Db 1381 SGLSRMLNDPPFQSYDEEVTWTSLELAHLKDFVSALPKDLHDECARGENLSVGQRLVCL 1440
QY 1441 ABALLRKTKILVLDATAAVDLETDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500
Db 1441 ABALLRKTKILVLDATAAVDLETDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500
QY 1501 DKGEIQEYCAPSDLLQORGLFYSMADAGLV 1531
Db 1501 DKGEIQEYCAPSDLLQORGLFYSMADAGLV 1531

```

RESULT 8

```

US-08-463-092B-2
; Sequence 2, Application US/08463092B
; Patent No. 5766880
; GENERAL INFORMATION:
; APPLICANT: Cole, Susan P.C.
; APPLICANT: Deeley, Roger G.
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ENCODING
; TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESS: PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS
; STREET: Queen's University at Kingston
; CITY: Kingston
; STATE: Ontario
; COUNTRY: CANADA
; ZIP: K7L 3N6
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,092B
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/966,923
; FILING DATE: 27-OCT-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/029,340
; FILING DATE: 8-MAR-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/141,893
; FILING DATE: 26-OCT-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/407,207
; FILING DATE: 20-MAR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Steeg, Carol Miernicki
; REGISTRATION NUMBER: 39,539
; REFERENCE/DOCKET NUMBER: Q1546
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (613) 545-2342
; TELEFAX: (613) 545-6853

```

```

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1531 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-463-092B-2

Query Match          79.3%; Score 7849; DB 1; Length 1531;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MALRGFCADGSDPLMDWNVTWNTSNPDFTKCFQNTVLVWVPCFVLMACFPFYLYLSRH 60
Db 1 MALRGFCADGSDPLMDWNVTWNTSNPDFTKCFQNTVLVWVPCFVLMACFPFYLYLSRH 60
QY 61 DRGIQMTPLNKTALGFLMIWCVADLFFYSFWESESGIFLAPVFLVSPILLGTTLLA 120
Db 61 DRGIQMTPLNKTALGFLMIWCVADLFFYSFWESESGIFLAPVFLVSPILLGTTLLA 120
QY 121 TFLIQLERRKGVQSSGIMLTFWLVLCALAILRSKIMTALKEDAQVDLFRDITFYVYFS 180
Db 121 TFLIQLERRKGVQSSGIMLTFWLVLCALAILRSKIMTALKEDAQVDLFRDITFYVYFS 180
QY 181 LLLIQLVLSCFSDRSPLFSETIHDNPNCPSSASPLSRITFWITGLIVGRQPLEGSD 240
Db 181 LLLIQLVLSCFSDRSPLFSETIHDNPNCPSSASPLSRITFWITGLIVGRQPLEGSD 240
QY 241 LWSLNKEDTSEQVVPVLVKNWKKCAKTRKQPKVYVSSKDPAPKESKVDANEVEAL 300
Db 241 LWSLNKEDTSEQVVPVLVKNWKKCAKTRKQPKVYVSSKDPAPKESKVDANEVEAL 300
QY 301 IVKSPQKEMNPSLPKVLKTYKTFPGYFLMSFFFKAIHDLMMFSGPQLKLLIKFVNDTKAPD 360
Db 301 IVKSPQKEMNPSLPKVLKTYKTFPGYFLMSFFFKAIHDLMMFSGPQLKLLIKFVNDTKAPD 360
QY 361 WQGYFTVLLFVTTACLTQLVLHQYPHICFVSGMRITKTAIVGAVRKALVITNSARKSSTV 420
Db 361 WQGYFTVLLFVTTACLTQLVLHQYPHICFVSGMRITKTAIVGAVRKALVITNSARKSSTV 420
QY 421 GEIVNLSVDAQRFMDLATYINMIWSAPLOVTLALYLWNLGSPVLAGVAVMVLMVFN 480
Db 421 GEIVNLSVDAQRFMDLATYINMIWSAPLOVTLALYLWNLGSPVLAGVAVMVLMVFN 480
QY 481 AVMAKTKTYQVAHMKSKDNRIKLANEILNGIKVLKYLAWELAFKDKVLAIQEEELKVLK 540
Db 481 AVMAKTKTYQVAHMKSKDNRIKLANEILNGIKVLKYLAWELAFKDKVLAIQEEELKVLK 540
QY 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTTIDENNILDAQTAFAVSLALFNILRFPNLI 600
Db 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTTIDENNILDAQTAFAVSLALFNILRFPNLI 600
QY 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWASDPPT 660
Db 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWASDPPT 660
QY 661 LMGITFISPEGALVAVVGQVCGKSSLLSALLAEMDKVEGHVAIKGSVAVYVQQAWIQND 720
Db 661 LMGITFISPEGALVAVVGQVCGKSSLLSALLAEMDKVEGHVAIKGSVAVYVQQAWIQND 720
QY 721 SLRENILFGCQLEEFYRSVIOACALLPDLPLPSGDRTEIGEKGVLNLSGGQKQVSLAR 780
Db 721 SLRENILFGCQLEEFYRSVIOACALLPDLPLPSGDRTEIGEKGVLNLSGGQKQVSLAR 780
QY 781 AVYSNADYILFDDPLSADVAHVGRKHFENVIGPKGMLKNKTRILYTHSHSYLPQVDVILV 840
Db 781 AVYSNADYILFDDPLSADVAHVGRKHFENVIGPKGMLKNKTRILYTHSHSYLPQVDVILV 840
QY 841 MSGGKISEMGSYQELLARDGAFABFLRTYASTEQSDAEENGVTGVSFGPKAKOMENGM 900
Db 841 MSGGKISEMGSYQELLARDGAFABFLRTYASTEQSDAEENGVTGVSFGPKAKOMENGM 900
QY 901 LVTDSAGKOLQRLSSSSSSSGDI1SRHHNSTAELQKAEAKBETWKLMEADKAQTQGVKL 960

```

Db 901 LVTDGAKQLQRLSSSSSSGDISRHNSTAELOKAEAKETWKLMEADRAQIGQVKL 960  
Qy 961 SVYDMYKAGLFIPLFSLFPMCHVHSALASNYLWLTDDPIVNGTOEHTKVLRSYVG 1020  
Db 961 SVYDMYKAGLFIPLFSLFPMCHVHSALASNYLWLTDDPIVNGTOEHTKVLRSYVG 1020  
Qy 1021 ALGISQIAGVGYMAVSGGILASRCLHVDLHLSILSRPMSFFERTSGNLVNRFSKEL 1080  
Db 1021 ALGISQIAGVGYMAVSGGILASRCLHVDLHLSILSRPMSFFERTSGNLVNRFSKEL 1080  
Qy 1081 DTWDSMIPVIXKFMGSLFNVIACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQL 1140  
Db 1081 DTWDSMIPVIXKFMGSLFNVIACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQL 1140  
Qy 1141 KELESVSRSPVYSHNETLLGVSVIRAEEOQRFHQSDLKVDENKQKAYPSIVANRWLA 1200  
Db 1141 KELESVSRSPVYSHNETLLGVSVIRAEEOQRFHQSDLKVDENKQKAYPSIVANRWLA 1200  
Qy 1201 VRLECVGNCIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTYLNLVMSSEMETNIVA 1260  
Db 1201 VRLECVGNCIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTYLNLVMSSEMETNIVA 1260  
Qy 1261 VERLKEYSETEKAPWQIQETAPPSSWPQVGRVFRNRYCLRYREDLDFVLRHINVTINGG 1320  
Db 1261 VERLKEYSETEKAPWQIQETAPPSSWPQVGRVFRNRYCLRYREDLDFVLRHINVTINGG 1320  
Qy 1321 EKVGVIGRTGAGKSSLTGLFRINSEAGEIIIDGINAKIGHDLRPKITIIPODPVLF 1380  
Db 1321 EKVGVIGRTGAGKSSLTGLFRINSEAGEIIIDGINAKIGHDLRPKITIIPODPVLF 1380  
Qy 1381 SGLSRMNDPFPQYDDEEVTLSLELAHLKDFVSALPDKLDHECAGGENLSVGORQLVCL 1440  
Db 1381 SGLSRMNDPFPQYDDEEVTLSLELAHLKDFVSALPDKLDHECAGGENLSVGORQLVCL 1440  
Qy 1441 ARALLRKTILVDBATAVLETDLQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
Db 1441 ARALLRKTILVDBATAVLETDLQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
Qy 1501 DKGEIQEYCAPSLLQORGLFYSMKADAGLV 1531  
Db 1501 DKGEIQEYCAPSLLQORGLFYSMKADAGLV 1531

RESULT 9

US-08-462-109A-2  
; Sequence 2, Application US/08462109A  
; Patent No. 582875  
; GENERAL INFORMATION:  
; APPLICANT: Cole, Susan P.C.  
; APPLICANT: Deeley, Roger G.  
; TITLE OF INVENTION: METHODS FOR IDENTIFYING  
; TITLE OF INVENTION: MULTIDRUG RESISTANT TUMOR CELLS  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: LAHIVE & COCKFIELD  
; STREET: 60 State Street, suite 510  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/462.109A  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/966, 923  
; FILING DATE: 27-OCT-1992

; APPLICATION NUMBER: 08/029,340  
; FILING DATE: 8-MAR-1993  
; APPLICATION NUMBER: 08/141,893  
; FILING DATE: 26-OCT-1993  
; APPLICATION NUMBER: 08/407,207  
; FILING DATE: 20-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: DeConti, Giulio A. Jr.  
; REGISTRATION NUMBER: 31,503  
; REFERENCE/DOCKET NUMBER: PQI-002CP4  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 227-5941  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1531 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-462-109A-2

Query Match 79.3%; Score 7849; DB 2; Length 1531;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MALRGFCSADGSDPLMDWNVTWNTSNPDFTKCFQNTVLVWVPCFVLWACFPFYLILSRH 60  
Db 1 MALRGFCSADGSDPLMDWNVTWNTSNPDFTKCFQNTVLVWVPCFVLWACFPFYLILSRH 60  
Qy 61 DRGVIQMTPLNKTALGFLWIVCWADLFYSFWRSGRIFLAPVFLVSPDLLGITTLA 120  
Db 61 DRGVIQMTPLNKTALGFLWIVCWADLFYSFWRSGRIFLAPVFLVSPDLLGITTLA 120  
Qy 121 TFLIQLERRKGVQSSGIMLTFWLVALCALAILRSKIMTALKEDAQVDLFRDITFYVYFS 180  
Db 121 TFLIQLERRKGVQSSGIMLTFWLVALCALAILRSKIMTALKEDAQVDLFRDITFYVYFS 180  
Qy 181 LLLIQLVLSFSDRSPLFSETHDPNCPSSASFLSRITFWWITGLIVRGVROPLEGSD 240  
Db 181 LLLIQLVLSFSDRSPLFSETHDPNCPSSASFLSRITFWWITGLIVRGVROPLEGSD 240  
Qy 241 LWSLNKEDTSEQVVPVLVQWKKCAKTRQPKVQVYSSKDPAPQKESKVDANEVEAL 300  
Db 241 LWSLNKEDTSEQVVPVLVQWKKCAKTRQPKVQVYSSKDPAPQKESKVDANEVEAL 300  
Qy 301 IVKSPQKEWNPSPFLKYLKTFPGPYFLMSFFPKAHLDMFSGPQILKLIKFNVDTKAPD 360  
Db 301 IVKSPQKEWNPSPFLKYLKTFPGPYFLMSFFPKAHLDMFSGPQILKLIKFNVDTKAPD 360  
Qy 361 WQGYFYTVLLFVTACLOTVLHQQYFHCYFVSGMRITAVIGAVYRKALVITNSARKSTV 420  
Db 361 WQGYFYTVLLFVTACLOTVLHQQYFHCYFVSGMRITAVIGAVYRKALVITNSARKSTV 420  
Qy 421 GEIVNLSVDAQRFMDLATYINMIWSAPLOQVILALYLLNLGFSVLGAVMVLWVPVN 480  
Db 421 GEIVNLSVDAQRFMDLATYINMIWSAPLOQVILALYLLNLGFSVLGAVMVLWVPVN 480  
Qy 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKLYAMELAFKDKVLAIROBELKVLK 540  
Db 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKLYAMELAFKDKVLAIROBELKVLK 540  
Qy 541 KSAVLSAVGTFTWVCTPFLVALCTFAVVVTIDENNILDAOTAFVLSALFNLIRPLNLTP 600  
Db 541 KSAVLSAVGTFTWVCTPFLVALCTFAVVVTIDENNILDAOTAFVLSALFNLIRPLNLTP 600  
Qy 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERPVDKGGGTNSITVRNATFTWARSDDPT 660  
Db 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERPVDKGGGTNSITVRNATFTWARSDDPT 660  
Qy 661 LINGITFSIPEGALVAVVGVQVCGKSLLSALLAEMDKVEGHVAIKGSVAVYPOQAWQND 720  
Db 661 LINGITFSIPEGALVAVVGVQVCGKSLLSALLAEMDKVEGHVAIKGSVAVYPOQAWQND 720

```

QY 721 SIRENIFGQLEEEPYRSVIOACALLPDLPLPGLDTEIGEKGWNLGGQKQVSLAR 780
Db 721 SIRENIFGQLEEEPYRSVIOACALLPDLPLPGLDTEIGEKGWNLGGQKQVSLAR 780
QY 781 AYSNADIYLPDPLSAVDHVGKHFENVIKPKMLKNTKILVTHSMYSYLPQVDVILV 840
Db 781 AYSNADIYLPDPLSAVDHVGKHFENVIKPKMLKNTKILVTHSMYSYLPQVDVILV 840
QY 841 MSGGKISEMGSOELLARDAFAELRTYASTEQEODAEENGVTGSGPGKEAKOMENGM 900
Db 841 MSGGKISEMGSOELLARDAFAELRTYASTEQEODAEENGVTGSGPGKEAKOMENGM 900
QY 901 LVTDSAGKOLQOLSSSSSYSGDI SRHNSIAELKAEAKETWKLMEADKAQGTQVKL 960
Db 901 LVTDSAGKOLQOLSSSSSYSGDI SRHNSIAELKAEAKETWKLMEADKAQGTQVKL 960
QY 961 SYVDYMKAGLFIPLSTFLPMCHVHSALASNYWLSLWTDPIVNGTQEHTKVLSVYG 1020
Db 961 SYVDYMKAGLFIPLSTFLPMCHVHSALASNYWLSLWTDPIVNGTQEHTKVLSVYG 1020
QY 1021 ALGISQGIAGVGYNAVSIGGILASRCLHVDLLHSLRSPMSFFERTPSGNLVNRFSKEL 1080
Db 1021 ALGISQGIAGVGYNAVSIGGILASRCLHVDLLHSLRSPMSFFERTPSGNLVNRFSKEL 1080
QY 1081 DTVDSMIPEVIRKPMGSLPNVICACIVILLATPIAAIIPPLGLIYFFVQRFYVASSRQL 1140
Db 1081 DTVDSMIPEVIRKPMGSLPNVICACIVILLATPIAAIIPPLGLIYFFVQRFYVASSRQL 1140
QY 1141 KLESVSRSPVYSHNETLLGVSVIRAPBEQERFIHQSDLKVDENKQKAYPSIVANRWLA 1200
Db 1141 KLESVSRSPVYSHNETLLGVSVIRAPBEQERFIHQSDLKVDENKQKAYPSIVANRWLA 1200
QY 1201 VRLCEGNCIVLFAALFAVISRHSLSAGLVGSVSLQVTTYLNLVMSSEMETNIVA 1260
Db 1201 VRLCEGNCIVLFAALFAVISRHSLSAGLVGSVSLQVTTYLNLVMSSEMETNIVA 1260
QY 1261 VERLKEYSETEAPWQIETAPPSSWPQVGRVFRNYCLRYREDLDFVLRHINVTINGG 1320
Db 1261 VERLKEYSETEAPWQIETAPPSSWPQVGRVFRNYCLRYREDLDFVLRHINVTINGG 1320
QY 1321 EKVIGVTRTGAKSSLTGLFRINSAEGEIIIDGINAKIGHDLRKFITIIPODPVLF 1380
Db 1321 EKVIGVTRTGAKSSLTGLFRINSAEGEIIIDGINAKIGHDLRKFITIIPODPVLF 1380
QY 1381 SGLSRNLDPFQSDDEEVTLSLELAHLKDFVSALPDKLDECAEGENLSVGQRLVCL 1440
Db 1381 SGLSRNLDPFQSDDEEVTLSLELAHLKDFVSALPDKLDECAEGENLSVGQRLVCL 1440
QY 1441 ARALARLTKILVLEATAAVDLETDLLIOSTIRTQFEDCTVLTIAHRLNTIMDYTRVIVL 1500
Db 1441 ARALARLTKILVLEATAAVDLETDLLIOSTIRTQFEDCTVLTIAHRLNTIMDYTRVIVL 1500
QY 1501 DKGEIQEYGAPSDLLOQRLGYFSMAKDAGLV 1531
Db 1501 DKGEIQEYGAPSDLLOQRLGYFSMAKDAGLV 1531

```

RESULT 10  
 US-08-460-907B-2  
 ; Sequence 2, Application US/08460907B  
 ; Patent No. 5891724

; GENERAL INFORMATION:  
 ; APPLICANT: Deeley, Roger G.  
 ; APPLICANT: Cole, Susan P.C.  
 ; TITLE OF INVENTION: METHODS FOR CONFERRING MULTIDRUG  
 ; TITLE OF INVENTION: RESISTANCE ON A CELL  
 ; NUMBER OF SEQUENCES: 9  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS  
 ; STREET: Queen's University at Kingston  
 ; CITY: Kingston  
 ; STATE: Ontario  
 ; COUNTRY: CANADA

```

; ZIP: K7L 3N6
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460.907B
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/966.923
; FILING DATE: 27-OCT-1992
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/029.340
; FILING DATE: 8-MAR-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/141.893
; FILING DATE: 26-OCT-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/407.207
; FILING DATE: 20-MAR-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Steeg, Carol Miernicki
; REGISTRATION NUMBER: 39,539
; REFERENCE/DOCKET NUMBER: Q1551
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (613) 545-2342
; TELEFAX: (613) 545-6853
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1531 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-460-907B-2

```

Query Match 79.3%; Score 7849; DB 2; Length 1531;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 MALRGFCSADGSDPLMDWNTWNTSNPDFTKCFQNTVLVWVPCFYLWACFPFYLYLSRH 60
Db 1 MALRGFCSADGSDPLMDWNTWNTSNPDFTKCFQNTVLVWVPCFYLWACFPFYLYLSRH 60
QY 61 DRGYIQTMTPLNKTKTALGFLMWIVCWADLFYFWSRSGIFLAPVFLVSPITLLGITTLA 120
Db 61 DRGYIQTMTPLNKTKTALGFLMWIVCWADLFYFWSRSGIFLAPVFLVSPITLLGITTLA 120
QY 121 TFLIQLERRKGVQSSGIMLTFWLVALCALAILRSKIMTALKEDAQVDLFRDITFYVYPS 180
Db 121 TFLIQLERRKGVQSSGIMLTFWLVALCALAILRSKIMTALKEDAQVDLFRDITFYVYPS 180
QY 181 LLLIQLVLSCSFSDRSPLESETIHDNPNCPSSASFLSRITFWITGLIVRGVROPLEGSD 240
Db 181 LLLIQLVLSCSFSDRSPLESETIHDNPNCPSSASFLSRITFWITGLIVRGVROPLEGSD 240
QY 241 LWSLNKEDTSEQVPEVLVKNWKECAKTRKOPKVVVYSSKDPAPQPKSSKVDANEVEAL 300
Db 241 LWSLNKEDTSEQVPEVLVKNWKECAKTRKOPKVVVYSSKDPAPQPKSSKVDANEVEAL 300
QY 301 IVKSPQKEWNPFLFKVLYKTFPGYFLMSPFFPKAIHDLMMFSGPQLKLLIKFVNQTKAPD 360
Db 301 IVKSPQKEWNPFLFKVLYKTFPGYFLMSPFFPKAIHDLMMFSGPQLKLLIKFVNQTKAPD 360
QY 361 WQGYEYTVLLFVTAQLOTIVLHQYFHCIVSGMRKTAIVIGAVYKALVITNSARKSSTV 420
Db 361 WQGYEYTVLLFVTAQLOTIVLHQYFHCIVSGMRKTAIVIGAVYKALVITNSARKSSTV 420

```

QY 421 GEIVNLSVDAQRFMDLATYINMIWSAPLQVILALYLLWNLGSPVLGAVVWLVMPVN 480  
DB 421 GEIVNLSVDAQRFMDLATYINMIWSAPLQVILALYLLWNLGSPVLGAVVWLVMPVN 480  
QY 481 AVMAKTKTYQVAHMKSKDNRIKLNWEILNGIKVLYAWELAFKQKVLAIROBELKVLK 540  
DB 481 AVMAKTKTYQVAHMKSKDNRIKLNWEILNGIKVLYAWELAFKQKVLAIROBELKVLK 540  
QY 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTTDENNILDQAQFVSLALFNILRFPNLILP 600  
DB 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTTDENNILDQAQFVSLALFNILRFPNLILP 600  
QY 601 MYISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGSGTNSITVRNATFTWASDDPT 660  
DB 601 MYISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGSGTNSITVRNATFTWASDDPT 660  
QY 661 LNGITFSIPEGALVAVVGQCGKSLLSALLAEMDKVEGHVAIKGSVAVYVPOQAWIQND 720  
DB 661 LNGITFSIPEGALVAVVGQCGKSLLSALLAEMDKVEGHVAIKGSVAVYVPOQAWIQND 720  
QY 721 SURENILFGCOLPEEYPSVIOACALLPDLPLPSGDRTEIGKGVNLGGOKORVSLAR 780  
DB 721 SURENILFGCOLPEEYPSVIOACALLPDLPLPSGDRTEIGKGVNLGGOKORVSLAR 780  
QY 781 AVYSNADIYLFDDPLSADVAHVGKHFENVIGPKGMLKXKTRILVTHSMSYLPQVDVIV 840  
DB 781 AVYSNADIYLFDDPLSADVAHVGKHFENVIGPKGMLKXKTRILVTHSMSYLPQVDVIV 840  
QY 841 MSGGKISEMGVQOELLARDGAPAEFLRTYVASTEQDAEENGVTGSGGKEAKOMENGM 900  
DB 841 MSGGKISEMGVQOELLARDGAPAEFLRTYVASTEQDAEENGVTGSGGKEAKOMENGM 900  
QY 901 LVTDGAGLOLQOLSSSSYSYSDISRHNSTAELOKAEAKKEETWKLMEADKAQGVKL 960  
DB 901 LVTDGAGLOLQOLSSSSYSYSDISRHNSTAELOKAEAKKEETWKLMEADKAQGVKL 960  
QY 961 SYVDMYKAIGLFIPLSIFLPMCNHVSALASNYWLSLWTDPIVNGTQEHKVLRSVYG 1020  
DB 961 SYVDMYKAIGLFIPLSIFLPMCNHVSALASNYWLSLWTDPIVNGTQEHKVLRSVYG 1020  
QY 1021 ALGISGIAVFGYSMAVSIIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFSEL 1080  
DB 1021 ALGISGIAVFGYSMAVSIIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFSEL 1080  
QY 1081 DTVDSMIEVIVKMGVSLFNVIACIVILLATPIAIIIPPLGLIYFFVQRFYVASSROL 1140  
DB 1081 DTVDSMIEVIVKMGVSLFNVIACIVILLATPIAIIIPPLGLIYFFVQRFYVASSROL 1140  
QY 1141 KELESVRSVPVSHFNETLLGVSIVIRAFEEQERFIHQSDLKVDENQKAYIPSIVANRWLA 1200  
DB 1141 KELESVRSVPVSHFNETLLGVSIVIRAFEEQERFIHQSDLKVDENQKAYIPSIVANRWLA 1200  
QY 1201 VRLCEVGNICVLFAALFAVISRHSISAGLVGLSVSYSLQVTTYLNLWVRMSEMETNIVA 1260  
DB 1201 VRLCEVGNICVLFAALFAVISRHSISAGLVGLSVSYSLQVTTYLNLWVRMSEMETNIVA 1260  
QY 1261 VERLKEYSETEKEAPWQOETAPPSWPQVGRVEFRNYCLRYREDLDFVLRHINTYINGG 1320  
DB 1261 VERLKEYSETEKEAPWQOETAPPSWPQVGRVEFRNYCLRYREDLDFVLRHINTYINGG 1320  
QY 1321 EVKGVGRTGAGKSSITLGLFRINSAEGEIIIDGINIAKIGLHDLRFPKTIIPDPVLF 1380  
DB 1321 EVKGVGRTGAGKSSITLGLFRINSAEGEIIIDGINIAKIGLHDLRFPKTIIPDPVLF 1380  
QY 1381 SGSLRMNLDPFQSYDEEVTWTSLELAHLKDFVSALPDKLDHECAGGENLSVGQQLVCL 1440  
DB 1381 SGSLRMNLDPFQSYDEEVTWTSLELAHLKDFVSALPDKLDHECAGGENLSVGQQLVCL 1440  
QY 1441 ABALLRKTILVLDATAVLDETDLQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
DB 1441 ABALLRKTILVLDATAVLDETDLQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
QY 1501 DKGEIQEYGAPSDLLQORGLFYSMADAGLV 1531

DB 1501 DKGEIQEYGAPSDLLQORGLFYSMADAGLV 1531

## RESULT 11

US-08-463-179A-2  
; Sequence 2, Application US/08463179A  
; Patent No. 6001563  
; GENERAL INFORMATION:  
; APPLICANT: Cole, Susan P.C.  
; APPLICANT: Deesley, Roger G.  
; TITLE OF INVENTION: METHODS FOR IDENTIFYING CHEMOSENSITIZERS  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street, suite 510  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/463.179A  
; FILING DATE:  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/966,923  
; FILING DATE: 27-OCT-1992  
; APPLICATION NUMBER: 08/029,340  
; FILING DATE: 8-MAR-1993  
; APPLICATION NUMBER: 08/141,893  
; FILING DATE: 26-OCT-1993  
; APPLICATION NUMBER: 08/407,207  
; FILING DATE: 20-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: DeConti, Giulio A. Jr.  
; REGISTRATION NUMBER: 31,503  
; REFERENCE/POCKET NUMBER: PQ1-002CF8  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 227-5941  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1531 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-463-179A-2

Query Match 79.3%; Score 7849; DB 3; Length 1531;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MALRGFCADSGDPLDWNVTWNTSNPDFTKCFONTVLVWVPCFYLWACFPFFFLYLSRH 60  
DB 1 MALRGFCADSGDPLDWNVTWNTSNPDFTKCFONTVLVWVPCFYLWACFPFFFLYLSRH 60  
QY 61 DRGIQMTPLNKTALGFLWVWADLFYSFWSRSGIFLAPVFLVSPDLLGITLLA 120  
DB 61 DRGIQMTPLNKTALGFLWVWADLFYSFWSRSGIFLAPVFLVSPDLLGITLLA 120  
QY 121 TFLIQLERRKGVQSSGIMLTFWLVALVCALAILRSKIMTALKEDAQVDFRDIITYVYFS 180  
DB 121 TFLIQLERRKGVQSSGIMLTFWLVALVCALAILRSKIMTALKEDAQVDFRDIITYVYFS 180  
QY 181 LLLQLVLSCFSDRSPLFSETIHDNPNCPSSASFLSRITFWITGLIVRGVROPLEGSD 240  
DB 181 LLLQLVLSCFSDRSPLFSETIHDNPNCPSSASFLSRITFWITGLIVRGVROPLEGSD 240

QY 241 LMSLNKEDTSEQVVPVLVQNMKKSCAKTRKQPVVYYSKDPAPKESKVDANEVEAL 300  
DB 241 LMSLNKEDTSEQVVPVLVQNMKKSCAKTRKQPVVYYSKDPAPKESKVDANEVEAL 300  
QY 301 IVKSPQKWNPSLKVLYKTFPGYELMSPFFKAIDHLMFSGPOTILKLLIKFVNDTKAPD 360  
DB 301 IVKSPQKWNPSLKVLYKTFPGYELMSPFFKAIDHLMFSGPOTILKLLIKFVNDTKAPD 360  
QY 361 MQGYFYTVLLFVTTACIQTLVLHQYFHI CFVSGMRKTAIVAGVYRKALVITNSARKSSTV 420  
DB 361 MQGYFYTVLLFVTTACIQTLVLHQYFHI CFVSGMRKTAIVAGVYRKALVITNSARKSSTV 420  
QY 421 GIVNLMMSVDAQRFMDLATYINMIWSAPLOVTLALYLLNLNLPSPVLAGVAVMLMPVN 480  
DB 421 GIVNLMMSVDAQRFMDLATYINMIWSAPLOVTLALYLLNLNLPSPVLAGVAVMLMPVN 480  
QY 481 AVAMKTKTYQVAHMKSKDNRIKLANEILNGIKVLYAWELA PKDKVLAIRQBELKVLK 540  
DB 481 AVAMKTKTYQVAHMKSKDNRIKLANEILNGIKVLYAWELA PKDKVLAIRQBELKVLK 540  
QY 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTTIDENNILDAQTAFAVSLALFNILRFPNLILP 600  
DB 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTTIDENNILDAQTAFAVSLALFNILRFPNLILP 600  
QY 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWARSDDPT 660  
DB 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWARSDDPT 660  
QY 661 LNGITFSPGALVAVGVQGGKSLASALLAEMDKVEGHVAIKGSVAVYVPPQAWIQND 720  
DB 661 LNGITFSPGALVAVGVQGGKSLASALLAEMDKVEGHVAIKGSVAVYVPPQAWIQND 720  
QY 721 SURENILFGCQLEEEYYSVIOACALLPDLEILPSGDRTEIGEGKVNLSGGOKQVSLAR 780  
DB 721 SURENILFGCQLEEEYYSVIOACALLPDLEILPSGDRTEIGEGKVNLSGGOKQVSLAR 780  
QY 781 AVYSNADYILFDPLISAVDAHVGKHI FENVIGPKGMLKNKTRILVTHSMSYLPQVDVIIV 840  
DB 781 AVYSNADYILFDPLISAVDAHVGKHI FENVIGPKGMLKNKTRILVTHSMSYLPQVDVIIV 840  
QY 841 MSGGKISEMGSYOELLARDGAEFLRTYASTEQDABEENGVTGSGGKEAKOMENGM 900  
DB 841 MSGGKISEMGSYOELLARDGAEFLRTYASTEQDABEENGVTGSGGKEAKOMENGM 900  
QY 901 LVTDGAGKQLQRLSSSSSYSGDISRHNSSTAELQKAEAKKEETWKLMEADKAQTGVK 960  
DB 901 LVTDGAGKQLQRLSSSSSYSGDISRHNSSTAELQKAEAKKEETWKLMEADKAQTGVK 960  
QY 961 SVYDYMKAIGLFIISFLIFPMCNHVSALASNYWLSLWTDPIVNGTQHTKVLSVYG 1020  
DB 961 SVYDYMKAIGLFIISFLIFPMCNHVSALASNYWLSLWTDPIVNGTQHTKVLSVYG 1020  
QY 1021 ALGISQGTAVFGYSMAVSTGGTILASRCLHVDLLHSILRSPMSFFERTSGNINVRFSKEL 1080  
DB 1021 ALGISQGTAVFGYSMAVSTGGTILASRCLHVDLLHSILRSPMSFFERTSGNINVRFSKEL 1080  
QY 1081 DTVDSDIMEVINKFMGSLFNIVGACIVILLATPIAAIIIPPLGLIYFFVQFPYVASSRQL 1140  
DB 1081 DTVDSDIMEVINKFMGSLFNIVGACIVILLATPIAAIIIPPLGLIYFFVQFPYVASSRQL 1140  
QY 1141 KRLESVSRSPVYSHFNETLLGVSVIRAFEEQBRFIHQSDLKVDENQKAYPSIVANRWLA 1200  
DB 1141 KRLESVSRSPVYSHFNETLLGVSVIRAFEEQBRFIHQSDLKVDENQKAYPSIVANRWLA 1200  
QY 1201 VRLCEVGNICIVLFAALFAVISHSLSAGLVGSYSLOVTTYLNLVIRMSSEMETNIVA 1260  
DB 1201 VRLCEVGNICIVLFAALFAVISHSLSAGLVGSYSLOVTTYLNLVIRMSSEMETNIVA 1260  
QY 1261 VERLKEYSETEKAPQIOETAPPSSWQVGVREFRNCLRYREDLDLFLVLRHINVTINGG 1320  
DB 1261 VERLKEYSETEKAPQIOETAPPSSWQVGVREFRNCLRYREDLDLFLVLRHINVTINGG 1320  
QY 1321 EKVGVIGRTGAGKSSLTGLFRINESABGEIIDIINIAGIKGLHDLRPFKTIIPQDPVLF 1380

DB 1321 EKVGVIGRTGAGKSSLTGLFRINESABGEIIDIINIAGIKGLHDLRPFKTIIPQDPVLF 1380  
QY 1381 SSGSRMNLDPFSQYSDEEVTSLSLAHLKDFVSALPDKLDHECAGGNNLSVGQRQLVCL 1440  
DB 1381 SSGSRMNLDPFSQYSDEEVTSLSLAHLKDFVSALPDKLDHECAGGNNLSVGQRQLVCL 1440  
QY 1441 ARALLRKTILVLEATAAVALDLEDDLIQSTIRTFQFEDCTVLTIAHRLNTIMDYTRVIL 1500  
DB 1441 ARALLRKTILVLEATAAVALDLEDDLIQSTIRTFQFEDCTVLTIAHRLNTIMDYTRVIL 1500  
QY 1501 DKGSIQFYCAPSDLLQORGLFYSMAKDAGLV 1531  
DB 1501 DKGSIQFYCAPSDLLQORGLFYSMAKDAGLV 1531

## RESULT 12

US-08-461-384B-2  
; Sequence 2, Application US/08461384B  
; Patent No. 6025473  
; GENERAL INFORMATION:  
; APPLICANT: Cole, Susan P.C.  
; APPLICANT: Deeley, Roger G.  
; TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS  
; CITY: Kingston  
; STATE: Ontario  
; COUNTRY: CANADA  
; ZIP: K7L 3N6  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/461,384B  
; FILING DATE: 05-JUN-95  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/966,923  
; FILING DATE: 27-OCT-1992  
; APPLICATION NUMBER: 08/029,340  
; FILING DATE: 8-MAR-1993  
; APPLICATION NUMBER: 08/141,893  
; FILING DATE: 26-OCT-1993  
; APPLICATION NUMBER: 08/407,207  
; FILING DATE: 20-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Steeg, Carol Miernicki  
; REGISTRATION NUMBER: 39,539  
; REFERENCE/DOCKET NUMBER: Q1547  
; TELEPHONE: (613) 545-2342  
; TELEFAX: (613) 545-6853  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1531 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-461-384B-2

Query Match 79.3%; Score 7849; DB 3; Length 1531;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MALRGFCSADGSDPLMDNVTWNTSNPDFTKCFQNTLVWVPCFYLMACFPFFYLYLSRH 60

DB 1 MALRGFCSADGSDPLMDNVTWNTSNPDFTKCFQNTLVWVPCFYLMACFPFFYLYLSRH 60

QY 61 DRGYQMTPLNKTALGFLWVCMADLFYSFWSRSRGI FLAPVFLVSPDLLGITLLA 120

Db 61 DRGIQWPLNKTALGFLWVCNADLFYSFERSRGIFLAPVFLVSPILLGITLLA 120  
Qy 121 TFLIQLERRKGQSSGIMLTFWLVALVCALAILRSKIMTALKEDAQVDFRDITFYVYFS 180  
Db 121 TFLIQLERRKGQSSGIMLTFWLVALVCALAILRSKIMTALKEDAQVDFRDITFYVYFS 180  
Qy 181 LLLIQLVLSRSDRSPLETHDNPDPSPASPLSRITFWITGLIVRGYRQPLEGSD 240  
Db 181 LLLIQLVLSRSDRSPLETHDNPDPSPASPLSRITFWITGLIVRGYRQPLEGSD 240  
Qy 241 LWSLNKEDTSQVVPVLVKNWKECAKTRKQPVVYSSKDPAPKESKVDANEEVEAL 300  
Db 241 LWSLNKEDTSQVVPVLVKNWKECAKTRKQPVVYSSKDPAPKESKVDANEEVEAL 300  
Qy 301 IVKSPQKEWNPSPLEKVLVKTGPFYLMGFFPKAIDHLMWFGSPQILKLLIKFVNDTKAPD 360  
Db 301 IVKSPQKEWNPSPLEKVLVKTGPFYLMGFFPKAIDHLMWFGSPQILKLLIKFVNDTKAPD 360  
Qy 361 WQGYFTYVLLFWTACLOTILVHOYPHICFVSGMRITKAVIGAVYRKALVITNSARKSTV 420  
Db 361 WQGYFTYVLLFWTACLOTILVHOYPHICFVSGMRITKAVIGAVYRKALVITNSARKSTV 420  
Qy 421 GEIVNLSVDAORFMDLATYINMIWSAPLOVTLALYLWNLGSPVLGAVVAVMLVMPVN 480  
Db 421 GEIVNLSVDAORFMDLATYINMIWSAPLOVTLALYLWNLGSPVLGAVVAVMLVMPVN 480  
Qy 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAWELAFKDKVLAIROBELKVLK 540  
Db 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAWELAFKDKVLAIROBELKVLK 540  
Qy 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTDENNILDAQAFVSLAFNLIRPLNLPL 600  
Db 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTDENNILDAQAFVSLAFNLIRPLNLPL 600  
Qy 601 WYISSIVQASVLSKELRIFLSHEELEPDSIERPVPKDGSGTNSITVRNATFTWARSDDPT 660  
Db 601 WYISSIVQASVLSKELRIFLSHEELEPDSIERPVPKDGSGTNSITVRNATFTWARSDDPT 660  
Qy 661 LINGITFSPEGALVAVGVQGVGCKSLLSALLAENDKVEGHVIAIKGSVAYVPPQAWIQND 720  
Db 661 LINGITFSPEGALVAVGVQGVGCKSLLSALLAENDKVEGHVIAIKGSVAYVPPQAWIQND 720  
Qy 721 SURENILFCQLEEPYRVSIVIOACALLPDLETLPSGDRTEIGEGKVLSSGGQKQVSLAR 780  
Db 721 SURENILFCQLEEPYRVSIVIOACALLPDLETLPSGDRTEIGEGKVLSSGGQKQVSLAR 780  
Qy 781 AVYSNADIYLFDDPLSAVDHVGKHFENVIGPKGMLKNKTRILVTHSMSYLPQVDVIV 840  
Db 781 AVYSNADIYLFDDPLSAVDHVGKHFENVIGPKGMLKNKTRILVTHSMSYLPQVDVIV 840  
Qy 841 MSGGKISEMGVQOELLARDGAFELRTYASTEQDAAEENGVTGSGPGEAKOMENG 900  
Db 841 MSGGKISEMGVQOELLARDGAFELRTYASTEQDAAEENGVTGSGPGEAKOMENG 900  
Qy 901 LVTDSAGKQLOQLSSSSSYSDISRHNSTAELOKAEAKETWKLMEADKAQGVKVL 960  
Db 901 LVTDSAGKQLOQLSSSSSYSDISRHNSTAELOKAEAKETWKLMEADKAQGVKVL 960  
Qy 961 SVYWDYMKAGIFIGPLSIFLPMCNHVSALASNYMLSLWTDPIVNGTQHTKVLRSVYG 1020  
Db 961 SVYWDYMKAGIFIGPLSIFLPMCNHVSALASNYMLSLWTDPIVNGTQHTKVLRSVYG 1020  
Qy 1021 ALGISOGIAVFGVSMVSTGGILASCLHVDLLHLSILRSPMSFFERTPSGNLVNPFSEL 1080  
Db 1021 ALGISOGIAVFGVSMVSTGGILASCLHVDLLHLSILRSPMSFFERTPSGNLVNPFSEL 1080  
Qy 1081 DTVDMSIPEVIMKMGSLFNIVGACIVILLATPIAAIIPPLGLIYFFVQRFYVASSRQL 1140  
Db 1081 DTVDMSIPEVIMKMGSLFNIVGACIVILLATPIAAIIPPLGLIYFFVQRFYVASSRQL 1140  
Qy 1141 KRLESVSRSPVYSHFNETLLGVSVIRAFERQERFIHQSDLKVDENOKAYYPSIVANRWLA 1200

Db 1141 KRLESVSRSPVYSHFNETLLGVSVIRAFERQERFIHQSDLKVDENOKAYYPSIVANRWLA 1200  
Qy 1201 VRLECVCNCIVLFAALFAVISRHSLSAGLVGLSVYSLOVTTYNLWLVNRMSSSEMTNIVA 1260  
Db 1201 VRLECVCNCIVLFAALFAVISRHSLSAGLVGLSVYSLOVTTYNLWLVNRMSSSEMTNIVA 1260  
Qy 1261 VERLKEYSETEKAPWIOETAPPSPWPQVGRVREFRNYCLRVEDLDFVLRHINVTINGG 1320  
Db 1261 VERLKEYSETEKAPWIOETAPPSPWPQVGRVREFRNYCLRVEDLDFVLRHINVTINGG 1320  
Qy 1321 EKVGIVGRTGAGKSSLTGLFRINESAEGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
Db 1321 EKVGIVGRTGAGKSSLTGLFRINESAEGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
Qy 1381 SGSLRMLNDPFSQYSDDEEVWTSLELAHLKDFVSGALPKLDHECAEGGENLSVGORQLVCL 1440  
Db 1381 SGSLRMLNDPFSQYSDDEEVWTSLELAHLKDFVSGALPKLDHECAEGGENLSVGORQLVCL 1440  
Qy 1441 ARALLRRTKILVLDEATAAVDLETDLLIQSTIRTQFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
Db 1441 ARALLRRTKILVLDEATAAVDLETDLLIQSTIRTQFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
Qy 1501 DKGEIQYAGAPSDLLQORGLFYSMADAGLV 1531  
Db 1501 DKGEIQYAGAPSDLLQORGLFYSMADAGLV 1531

## RESULT 13

US-08-407-207A-2

; Sequence 2, Application US/08407207A

; Patent No. 6063621

; GENERAL INFORMATION:

; APPLICANT: Deeley, Roger G.

; APPLICANT: Cole, Susan P.C.

; TITLE OF INVENTION: ANTIBODIES TO A MULTIDRUG RESISTANCE PROTEIN

; NUMBER OF SEQUENCES: 7

; CORRESPONDENCE ADDRESS:

; ADDRESS: PARTEQ RESEARCH &amp; DEVELOPMENT INNOVATIONS

; STREET: Queen's University at Kingston

; CITY: Kingston

; STATE: Ontario

; COUNTRY: CANADA

; ZIP: K7L 3N6

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: ASCII text

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/407,207A

; FILING DATE: 20-MAR-1995

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/966,923

; FILING DATE: 27-OCT-1992

; APPLICATION NUMBER: 08/029,340

; FILING DATE: 8-MAR-1993

; APPLICATION NUMBER: 08/141,893

; FILING DATE: 26-OCT-1993

; ATTORNEY/AGENT INFORMATION:

; NAME: Steeg, Carol Miernicki

; REGISTRATION NUMBER: 39,539

; REFERENCE/DOCKET NUMBER: Q1512

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (613) 545-2342

; TELEFAX: (613) 545-6853

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1531 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-407-207A-2

Query Match			
Best Local Similarity 79.3%; Score 7849; DB 3; Length 1531;			
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;			
QY	1	MALRGCSDGSDPLMDNNTWNTSNPDTKCFONTVLVWPCFYLWACFPFYLYLSRH	60
Db	1	MALRGCSDGSDPLMDNNTWNTSNPDTKCFONTVLVWPCFYLWACFPFYLYLSRH	60
QY	61	DRGYQMTPLNKTATGALGELLVWCADLIFYFWERSRGIFLAPFLVSPPTLLGTTLLA	120
Db	61	DRGYQMTPLNKTATGALGELLVWCADLIFYFWERSRGIFLAPFLVSPPTLLGTTLLA	120
QY	121	TELIOLERRKGVQSSGIMLTFWLVALCALAILRSKIMTALKEDAQVDLFRDITFYVYFS	180
Db	121	TELIOLERRKGVQSSGIMLTFWLVALCALAILRSKIMTALKEDAQVDLFRDITFYVYFS	180
QY	181	LALIQLVLSCFSDRSPLSETIHDNCPCESSASFLSRITFWITGLIVRGYPQLEGSD	240
Db	181	LALIQLVLSCFSDRSPLSETIHDNCPCESSASFLSRITFWITGLIVRGYPQLEGSD	240
QY	241	LWLNKEDTSEOVVPLVKNWKECAKTRKQPVVYSSKDPAPKESKVDANEVEAL	300
Db	241	LWLNKEDTSEOVVPLVKNWKECAKTRKQPVVYSSKDPAPKESKVDANEVEAL	300
QY	301	IVKSPQKWNPSLFKVLKTFPGPYFLMSFFFKAIHDLMMFSGPQILKLIKFVNDTKAPD	360
Db	301	IVKSPQKWNPSLFKVLKTFPGPYFLMSFFFKAIHDLMMFSGPQILKLIKFVNDTKAPD	360
QY	361	WQGYFTVLLFTVACLOTULVHOYFHI CFVSGMRKTAIVGAYRKALVITNSARKSTV	420
Db	361	WQGYFTVLLFTVACLOTULVHOYFHI CFVSGMRKTAIVGAYRKALVITNSARKSTV	420
QY	421	GEIVNLMVDAORFMDLATYINWISAPLOVILALYLWNLGSPSLAGVAVWMLMVPVN	480
Db	421	GEIVNLMVDAORFMDLATYINWISAPLOVILALYLWNLGSPSLAGVAVWMLMVPVN	480
QY	481	AVWAMKTQTYQVAHMKSDNRILKLNELINGIKVLKLYAWELAFKDKVLAIROBELKVLK	540
Db	481	AVWAMKTQTYQVAHMKSDNRILKLNELINGIKVLKLYAWELAFKDKVLAIROBELKVLK	540
QY	541	KSAYLSAVGTFTVCTPFLVALCTFAVYVITDENNILDQAOTAFVSLALFNILRFPNLTLP	600
Db	541	KSAYLSAVGTFTVCTPFLVALCTFAVYVITDENNILDQAOTAFVSLALFNILRFPNLTLP	600
QY	601	MWISSIVQASVSLKRLIRIFLSHEELEPDSIERPVKDGGTNSITVRNATFTWARSDDPT	660
Db	601	MWISSIVQASVSLKRLIRIFLSHEELEPDSIERPVKDGGTNSITVRNATFTWARSDDPT	660
QY	661	LNGITFISIPEGALVAVVGVQCGKSLLSALLAEMDKVEGHVAIKGSVAVYVPPQAWIQND	720
Db	661	LNGITFISIPEGALVAVVGVQCGKSLLSALLAEMDKVEGHVAIKGSVAVYVPPQAWIQND	720
QY	721	SURENILFCQLEEPYRSVIOACALLPDLPLSGDRTEIGEKNVLSGGQKQVSLAR	780
Db	721	SURENILFCQLEEPYRSVIOACALLPDLPLSGDRTEIGEKNVLSGGQKQVSLAR	780
QY	781	AVYSNADIYLFDDPLSADAHVGHKIFENVIGPKGMLKNKTRILVTHSMSYLPQVDVIV	840
Db	781	AVYSNADIYLFDDPLSADAHVGHKIFENVIGPKGMLKNKTRILVTHSMSYLPQVDVIV	840
QY	841	MSGGKISEMGSYQELIARDGAEFLRTYASTEQEQAENGVTGSGPGKEAKQWENG	900
Db	841	MSGGKISEMGSYQELIARDGAEFLRTYASTEQEQAENGVTGSGPGKEAKQWENG	900
QY	901	LVTDSAGKQQLQOLSSSSSYSGDISRHNSSTAELOKAEAKBETWKLMEADKAQTCQVKL	960
Db	901	LVTDSAGKQQLQOLSSSSSYSGDISRHNSSTAELOKAEAKBETWKLMEADKAQTCQVKL	960
QY	961	SVYWDYMKAGLGFISFLSIFLFCMCHVSAALSNYLSLWTDPIVNGTOEHTKVRLSVYG	1020
Db	961	SVYWDYMKAGLGFISFLSIFLFCMCHVSAALSNYLSLWTDPIVNGTOEHTKVRLSVYG	1020

RESULT 14

US-08-463-092B-6  
; Sequence 6, Application US/08463092B  
; Patent No. 5766880  
; GENERAL INFORMATION:  
; APPLICANT: Cole, Susan P.C.  
; APPLICANT: Deeley, Roger G.  
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ENCODING  
; TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS  
; STREET: Queen's University at Kingston  
; CITY: Kingston  
; STATE: Ontario  
; COUNTRY: CANADA  
; ZIP: K7L 3N6  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/463,092B  
; FILING DATE: 05-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/966,923  
; FILING DATE: 27-OCT-1992  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/029,340  
; FILING DATE: 8-MAR-1993  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/141,893

QY	720	DSLRENILFGQLEEPYRSVIOACALLPDLLEILPSGDRTEIGEKGYNLSGGQKQVSLA	779
DB	720	DSLRENILFGHPLQENYKAVMEACALLPDLLEILPSGDRTEIGEKGYNLSGGQKQVSLA	779
QY	780	RAVYSNADIYLFDDPLSADAHVGHKHFENVIGPKMLKVKTRILVTHSHSYLPQVDVII	839
DB	780	RAVYSNSDIYLFDDPLSADAHVGHKHFENVIGPKMLKVKTRILVTHSHSYLPQVDVII	839
QY	840	VMSGKISEMGYSOELLARDGAFELRTYASTEEOQDAEENGVTGSGPKAKOMENG	899
DB	840	VMSGKISEMGYSOELLARDGAFELRTYASTEEOQDAEENGVTGSGPKAKOMENG	899
QY	900	MLVTDGAKQLQRLSSSSSYSGDISRHHNSTAELQKAEKCBETWKLMEADKAQTOQVX	959
DB	900	MLVTDGAKQLQRLSSSSSYSGDISRHHNSTAELQKAEKCBETWKLMEADKAQTOQVX	959
QY	960	LSVTWYDKAIGLPIFSLFPMCNHVSALASNYWLSLWTD--PIVNGTQEHKTVLSV	1018
DB	960	LSVTWYDKAIGLPIFSLFPMCNHVSALASNYWLSLWTD--PIVNGTQEHKTVLSV	1018
QY	1019	YGALGISOGLAVFGYSMAVSIIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFPSK	1078
DB	1019	YGALGISOGLAVFGYSMAVSIIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFPSK	1078
QY	1079	ELDTVDSDMIPEVIMKFMGSLFNIVGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSR	1138
DB	1079	ELDTVDSDMIPEVIMKFMGSLFNIVGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSR	1138
QY	1139	QLKLESVSRSPVYSHFNELLGVSVIRAFEEQERFHQSGLKVDENOKAYPSIVANRW	1198
DB	1139	QLKLESVSRSPVYSHFNELLGVSVIRAFEEQERFHQSGLKVDENOKAYPSIVANRW	1198
QY	1199	LAVLECVGNCIVLFAALFAVISHSLSAGLVLSVSLQVTVTYLNLVMSSEMETNI	1258
DB	1199	LAVLECVGNCIVLFAALFAVISHSLSAGLVLSVSLQVTVTYLNLVMSSEMETNI	1258
QY	1259	VAVERLKEYSETEKEAPWQIQETAPPSTWPHSGRVEPRDYCLRYREDLVLKHINTIE	1315
DB	1259	VAVERLKEYSETEKEAPWQIQETAPPSTWPHSGRVEPRDYCLRYREDLVLKHINTIE	1315
QY	1319	GGEKVGIVRTGAKGSSITGLFRINESAGEIIIDGINTAKIGLHDLRFKTIIPQDPV	1378
DB	1319	GGEKVGIVRTGAKGSSITGLFRINESAGEIIIDGINTAKIGLHDLRFKTIIPQDPV	1378
QY	1379	LFSGSLRNLDPPFQYSDSEVWTSLELAHLKDFVSALPDKLDHECAEGENLSVGQRLV	1438
DB	1379	LFSGSLRNLDPPFQYSDSEVWTSLELAHLKDFVSALPDKLDHECAEGENLSVGQRLV	1438
QY	1439	CLARALLRKTILVLDDEATAADLETDLLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVI	1498
DB	1439	CLARALLRKTILVLDDEATAADLETDLLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVI	1498
QY	1499	VLDKGEIQEYCAPSDLLQORGLFYSMADAGLV 1531	
DB	1499	VLDKGEIQEYCAPSDLLQORGLFYSMADAGLV 1531	
QY	1496	VLDKGEVREGCAPSLLQORGLFYSMADAGLV 1528	
DB	1496	VLDKGEVREGCAPSLLQORGLFYSMADAGLV 1528	
RESULT 15			
US-08-462-109A-6			
; Sequence 6, Application US/08462109A			
; Patent No. 5882875			
; GENERAL INFORMATION:			
; APPLICANT: Cole, Susan P.C.			
; APPLICANT: Deely, Roger G.			
; TITLE OF INVENTION: METHODS FOR IDENTIFYING			
; TITLE OF INVENTION: MULTIDRUG RESISTANT TUMOR CELLS			
; NUMBER OF SEQUENCES: 6			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: LAHIVE & COCKFIELD			
; STREET: 60 State Street, suite 510			
; CITY: Boston			
; STATE: Massachusetts			
; COUNTRY: USA			

QY	1	MALRFGCSADGSDPLDNDWNTWNTSNPDFTKCFQNTVLVWVPCFYLWACEFPFYFLYLSRH	60
DB	1	MALRFGCSADGSDPLDNDWNTWNTSNPDFTKCFQNTVLVWVPCFYLWACEFPFYFLYLSRH	60
QY	61	DRGYQMTPLNKTALGFLMIVCWADLFYSFWRSGIFLAPVFLVPSLTLLGTTLLA	120
DB	61	DRGYQMTPLNKTALGFLMIVCWADLFYSFWRSGIFLAPVFLVPSLTLLGTTLLA	120
QY	121	TFLIOLERRKGVSIGIMTFMVLVALCALAILRSKIMTALKEDAQDLFRDITFYVFS	180
DB	121	TFLIOLERRKGVSIGIMTFMVLVALCALAILRSKIMTALKEDAQDLFRDITFYVFS	180
QY	181	LLLIOLVLSGFCSDRPLSETHTDNPCESSASFLSRTTFFWITGLIVRGVROPLEGSD	240
DB	181	LLLIOLVLSGFCSDRPLSETHTDNPCESSASFLSRTTFFWITGLIVRGVROPLEGSD	240
QY	241	LWSLNKEDTSEGVVPLVKNWKECAKTRKQPVVYIS-SKDPAPQKSSSKVDANEVEA	299
DB	241	LWSLNKEDTSEGVVPLVKNWKECAKTRKQPVVYIS-SKDPAPQKSSSKVDANEVEA	299
QY	300	LIVKSPQKWNLSFKVLYKTGPFYFLSPFKATHDLMFSGPOIILKLIKFNVDTKAP	359
DB	300	LIVKSPQKWNLSFKVLYKTGPFYFLSPFKATHDLMFSGPOIILKLIKFNVDTKAP	359
QY	360	DMQGYFYTVLLFVTAQTLVHLQYFHCIFVSGMRKTAIVAGVYRKALVITNSARKSST	419
DB	360	DMQGYFYTVLLFVTAQTLVHLQYFHCIFVSGMRKTAIVAGVYRKALVITNSARKSST	419
QY	420	VGEIVNLSVDAQRFMDLATYINMIWSAPLOVILALYLWMLGSPVLGAVWVLMVVPV	479
DB	420	VGEIVNLSVDAQRFMDLATYINMIWSAPLOVILALYLWMLGSPVLGAVWVLMVVPV	479
QY	480	NAMAMKTTYQVAHMKSDNRIKLWNEILNGIKVLYAWELAFKDKVLAIRQBELKVL	539
DB	480	NAMAMKTTYQVAHMKSDNRIKLWNEILNGIKVLYAWELAFKDKVLAIRQBELKVL	539
QY	540	KKSAYLSAVGTFTWCTPFLVALCFVAVVTTIDENNILDAQAFVSLALFNILRPLNLL	599
DB	540	KKSAYLSAVGTFTWCTPFLVALCFVAVVTTIDENNILDAQAFVSLALFNILRPLNLL	599
QY	600	PMVISSIVQASVSLKRLIFLSHEELEPDSIERRPVKDGGSNTSITVRNATFTWARSPP	659
DB	600	PMVISSIVQASVSLKRLIFLSHEELEPDSIERRPVKDGGSNTSITVRNATFTWARSPP	659
QY	660	TLNGITFTSPEGALVAVGVGCGKSSLLSALLAEMDKVEGHVAKGSAVYVPPQAWION	719
DB	660	TLNGITFTSPEGALVAVGVGCGKSSLLSALLAEMDKVEGHVAKGSAVYVPPQAWION	719

FILING DATE: 26-OCT-1993			
CLASSIFICATION: 435			
PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: 08/407,207			
; FILING DATE: 20-MAR-1995			
CLASSIFICATION: 435			
ATTORNEY/AGENT INFORMATION:			
; NAME: Steeg, Carol Miernicki			
; REGISTRATION NUMBER: 39,539			
; REFERENCE/DOCKET NUMBER: Q1546			
TELECOMMUNICATION INFORMATION:			
; TELEPHONE: (613) 545-6853			
; TELEFAX: (613) 545-6853			
; INFORMATION FOR SEQ ID NO. 6:			
SEQUENCE CHARACTERISTICS:			
; LENGTH: 1528 amino acids			
; TYPE: amino acid			
; TOPOLOGY: linear			
; MOLECULE TYPE: protein			
US-08-463-092B-6			
Query Match			
Best Local Similarity 80.7%; Score 7002.5; DB 1; Length 1528;			
Matches 1349; Conservative 102; Mismatches 75; Indels 7; Gaps 5;			

```

; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/462.109A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/966.923
; FILING DATE: 27-OCT-1992
; APPLICATION NUMBER: 08/029.340
; FILING DATE: 8-MAR-1993
; APPLICATION NUMBER: 08/141.893
; FILING DATE: 26-OCT-1993
; APPLICATION NUMBER: 08/407.207
; FILING DATE: 20-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: DeConti, Giulio A. Jr.
; REGISTRATION NUMBER: 31.503
; REFERENCE/DOCKET NUMBER: PQI-002CP4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1528 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-462-109A-6

```

```

Query Match 70.7%; Score 7002.5; DB 2; Length 1528;
Best Local Similarity 88.0%; Pred. No. 0;
Matches 1349; Conservative 102; Mismatches 75; Indels 7; Gaps 5;

QY 1 MALRFGCSADGSDPLWNNVTWNTSGNPDFTKCFQNTVLVWVPCFYLWACFPFYLYLSRH 60
DB 1 MALRFGCSADGSDPLWNNVTWNTSGNPDFTKCFQNTVLVWVPCFYLWACFPFYLYLSRH 60
QY 61 DRGYIQMTPLNKTALGFLWIVCWADLFYFWSRSRGIFLAPVFLVSPITLLGTTLLA 120
DB 61 DRGYIQMTPLNKTALGFLWIVCWADLFYFWSRSRGIFLAPVFLVSPITLLGTTLLA 120
QY 121 TFLQLERRKGVOSSGIMLTFWLVVALVCAIATLRKIMTALKEDAQVDLFRDITFWYFES 180
DB 121 TFLQLERRKGVOSSGIMLTFWLVVALVCAIATLRKIMTALKEDAQVDLFRDITFWYFES 180
QY 181 LLLIQLVLCFSDRPLFSETIHDNPPCESSASFLSRITFWITGLIVRGYRQPLEGSD 240
DB 181 LVLVQLVLCFSDRPLFSETIHDNPPCESSASFLSRITFWITGLIVRGYRQPLEGSD 240
QY 241 LWSLNKEDTSEQVVPVLVNNWKECAKTRQPKVYVS-SKDPAPQKSSKVDANEVEEA 299
DB 241 LWSLNKEDTSEQVVPVLVNNWKECAKTRQPKVYVS-SKDPAPQKSSKVDANEVEEA 300
QY 300 LIVSPQKWNNSFLKVLTKTGPFLMSFFPKAHDLMFSGPQILKLIKFNVDTKAP 359
DB 301 LIVSPHKBREPSFLKVLTKTGPFLMSFFPKAHDLMFSGPQILKLIKFNVDREAP 360
QY 360 DMQGYFTYLLVFTACQLVHLQHFPHICFVSGMRITKAVIGAVYRKALVITNSARKSST 419
DB 361 DMQGYFTYLLVFTACQLVHLQHFPHICFVSGMRITKAVIGAVYRKALVITNSARKSST 420
QY 420 VGEIVNLSVDAQRFMDLATYINMIWSAPLQVILALYLWNLGSPSVLAGVAVMLVMPV 479
DB 421 VGEIVNLSVDAQRFMDLATYINMIWSAPLQVILALYLWNLGSPSVLAGVAVMLVMPV 480
QY 480 NAVMAMKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAWELAFKDKVLAIROBELKVL 539
DB 481 NAVMAMKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAWELAFKDKVMISROBELKVL 540

```

```

QY 540 KKSAYLSAVGFTTWCCTPFLVALCTFAVYVTVIDENNILDQAOTAFVSLALFNILRPPNLI 599
DB 541 KKSAYLAAGVGTFTWCCTPFLVALSTFAVFTVDERNILDAKKAFVSLALFNILRPPNLI 600
QY 600 PMVISSIVQASVSLKRLRIFLSHEELPDSLERPVKDGCGTNSITVNRNFTTWARSDPP 659
DB 601 PMVISSIVQASVSLKRLRIFLSHEELPDSLERPVKDGCGTNSITVNRNFTTWARSDPP 659
QY 660 TLNGITFSIPGALVAVVGVCGCKSSLLSALLAEMDKVEGHVAIKGSVAVVPPQAWIQN 719
DB 660 TLNGITFSIPGALVAVVGVCGCKSSLLSALLAEMDKVEGHVTLKGSVAVVPPQAWIQN 719
QY 720 DSLRENILFGQLBEPYRSVIQCALLPDLEIILPSGDRTEIGEGVNLSGQKQKRVSLA 779
DB 720 DSLRENILFGHPLQENYKAVWEACALLPDLEIILPSGDRTEIGEGVNLSGQKQKRVSLA 779
QY 780 RAVYSNADIYLFDDPLSAVDHVGKHFENYVIGPKMLKNKTRILILVTHSMVLPDQVII 839
DB 780 RAVYSNSDIYLFDDPLSAVDHVGKHFENYVIGPKMLKNKTRILILVTHSGISYLPDQVII 839
QY 840 VMSGKISEMGSYQELLARDGAFAEFLRTYASTEQEQDAEENGVTGVSQPGKEAKOMENG 899
DB 840 VMSGKISEMGSYQELLARDGAFAEFLRTYASTEQEQDAEENGVTGVSQPGKEAKOMENG 899
QY 900 MLVTSAGKQLOROLSSSSSYSGDISRHNSHTAELOKABAKKEETWKLMEADKAQTGVK 959
DB 897 MLVTDTVGKHLQRLHLSNSSHSGDTSQOHSSIAELQKAGA-KEETWKLMEADKAQTGVK 955
QY 960 LSVYNDYMKATGLFISLSEIFLFCNVHVSALANVWLSLWTD--PUNGTQEHKTVLSV 1018
DB 956 LSVYNNYMKATGLFITLSLSEIFLFCNVHVSALANVWLSLWTD--PUNGTQEHKTVLSV 1015
QY 1019 YGALGISOGIAVFGYSMAVSIIGGILASRCLHVDLHLSILRSPMSPFFERTPSGNLNRFSK 1078
DB 1016 YGALGILGAAIFGYSMAVSIIGGIFASRLHLDLLVNLRSMPSEFFERTPSGNLNRFSK 1075
QY 1079 ELDTVDSMIPVIMKFMGSLFNVIACIVILLATPIAAIIIPPLGLIYFFVQRFVASSR 1138
DB 1076 ELDTVDSMIPVIMKFMGSLFVIGAVIIILLATPIAAVIPPPLGLVYFFVQRFVASSR 1135
QY 1139 QLKLESVRSRPVSHFNETLLGVSVTRAFSEQERFTHQSDIKVDENOKAYVPSIVANRW 1198
DB 1136 QLKLESVRSRPVSHFNETLLGVSVTRAFSEQERFTHQSDIKVDENOKAYVPSIVANRW 1195
QY 1199 LAVRLECVGNCIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTYLNMLVRMSSEMETNI 1258
DB 1196 LAVRLECVGNCIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTYLNMLVRMSSEMETNI 1255
QY 1259 VAVRLEKEYSETEKAPMQIOETAPPSPQVGRVFRNYCLURYREDLDFVLURHINVTIN 1318
DB 1256 VAVRLEKEYSETEKAPMQIOETAPPSPQVGRVFRNYCLURYREDLDFVLURHINVTIN 1315
QY 1319 GGEKGVIVGRTGAGKSSITLGLFRINESAEGEIIIDGINIAKIGLHDLRFKTIIPQDPV 1378
DB 1316 GGEKGVIVGRTGAGKSSITLGLFRINESAEGEIIIDGINIAKIGLHDLRFKTIIPQDPV 1375
QY 1379 LFGSLRNWLDPPFSQYSDSEEVMTSLAHLKDFVSALPKDHECAEGENLSVGORQIV 1438
DB 1376 LFGSLRNWLDPPFSQYSDSEEVMTSLAHLKDFVSALPKDHECAEGENLSVGORQIV 1435
QY 1439 CLARALLRKTILVLDEATAAVDLETDLIQSTIRTFQEDCTVLTIAHRLNTIMDYTRVI 1498
DB 1436 CLARALLRKTILVLDEATAAVDLETDLIQSTIRTFQEDCTVLTIAHRLNTIMDYTRVI 1495
QY 1499 VLDKGEIYEGAPSDLLQORGLFYFSMAKDAGLV 1531
DB 1496 VLDKGEVRECCAPSELLOQRGIFYSMAKDAGLV 1528

```

Search completed: March 18, 2005, 11:03:39  
Job time : 47.2909 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2005, 11:03:55 ; Search time 609.338 Seconds  
(without alignments)  
1045.320 Million cell updates/sec

Title: US-10-665-283-1

Perfect score: 9903

Sequence: 1 MALRGFCSADGSDPLMDWNV.....RSVAVAKPKFSIPDSLS 1927

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1401741 seqs, 330541175 residues

Total number of hits satisfying chosen parameters: 1401741

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	7860	79.4	1531	11	US-09-939-853A-86
2	7860	79.4	1531	16	US-10-667-891-6
3	7860	79.4	1531	17	US-10-484-577-678
4	7769	78.5	1515	11	US-09-939-853A-87
5	7037	71.1	1388	16	US-10-408-765A-1718
6	7002.5	70.7	1528	11	US-09-939-853A-88
7	4491.5	45.4	1527	11	US-09-939-853A-83
8	4491.5	45.4	1527	11	US-09-939-853A-84
9	4491.5	45.4	1527	15	US-10-295-027-1342
10	4396	44.4	1522	11	US-09-939-853A-85
11	3860.5	39.0	1548	16	US-10-667-891-1
12	3603.5	36.4	1505	16	US-10-667-891-4
13	3601.5	36.4	1499	16	US-10-667-891-3

14	3600	36.4	1542	15	US-10-363-112-4	Sequence 4, Appli
15	3600	36.4	1545	15	US-10-252-155-4	Sequence 4, Appli
16	3600	36.4	1545	15	US-10-252-155-24	Sequence 24, Appli
17	3600	36.4	1545	15	US-10-252-155-26	Sequence 26, Appli
18	3600	36.4	1545	15	US-10-252-155-28	Sequence 28, Appli
19	3600	36.4	1545	15	US-10-252-155-30	Sequence 30, Appli
20	3600	36.4	1545	15	US-10-252-155-32	Sequence 32, Appli
21	3600	36.4	1545	15	US-10-252-155-34	Sequence 34, Appli
22	3600	36.4	1545	15	US-10-252-155-38	Sequence 40, Appli
23	3600	36.4	1545	15	US-10-252-155-40	Sequence 40, Appli
24	3600	36.4	1545	15	US-10-252-155-46	Sequence 46, Appli
25	3600	36.4	1545	15	US-10-252-155-48	Sequence 48, Appli
26	3600	36.4	1545	15	US-10-363-112-2	Sequence 2, Appli
27	3600	36.4	1545	15	US-10-363-112-6	Sequence 6, Appli
28	3600	36.4	1545	15	US-10-363-112-8	Sequence 8, Appli
29	3600	36.4	1545	15	US-10-363-112-10	Sequence 10, Appli
30	3600	36.4	1545	15	US-10-363-112-12	Sequence 12, Appli
31	3600	36.4	1545	15	US-10-363-112-14	Sequence 14, Appli
32	3600	36.4	1545	15	US-10-363-112-16	Sequence 16, Appli
33	3599.5	36.3	1507	16	US-10-667-891-5	Sequence 5, Appli
34	3599	36.3	1545	15	US-10-252-155-44	Sequence 42, Appli
35	3597	36.3	1545	15	US-10-252-155-42	Sequence 42, Appli
36	3595	36.3	1545	16	US-10-741-601-369	Sequence 369, App
37	3595	36.3	1545	16	US-10-741-601-370	Sequence 370, App
38	3593	36.3	1545	15	US-10-252-155-36	Sequence 36, Appli
39	3582	36.2	1545	15	US-10-252-155-604	Sequence 604, App
40	3385.5	34.2	1503	10	US-09-792-616-3	Sequence 3, Appli
41	3385.5	34.2	1503	16	US-10-764-328-3	Sequence 4, Appli
42	3385.5	34.2	1503	17	US-10-741-600-1583	Sequence 1583, Ap
43	3289	33.2	1494	15	US-10-369-493-6697	Sequence 6697, Ap
44	3225	32.6	1573	15	US-10-369-493-6867	Sequence 6867, Ap
45	3225	32.6	1573	15	US-10-369-493-6868	Sequence 6868, Ap

ALIGNMENTS

RESULT 1  
US-09-939-853A-86  
; Sequence 86, Application US/09939853A  
; Publication No. US20040039163A1  
; GENERAL INFORMATION:  
; APPLICANT: Burgess et al.  
; TITLE OF INVENTION: No. US20040039163A1el Proteins and Nucleic Acids Encoding Same  
; FILE REFERENCE: 21402-099  
; CURRENT APPLICATION NUMBER: US/09/939,853A  
; CURRENT FILING DATE: 2001-08-27  
; PRIOR APPLICATION NUMBER: 60/228,191  
; PRIOR FILING DATE: 2000-08-25  
; PRIOR APPLICATION NUMBER: 60/267,300  
; PRIOR FILING DATE: 2001-02-08  
; PRIOR APPLICATION NUMBER: 60/269,961  
; PRIOR FILING DATE: 2001-02-20  
; PRIOR APPLICATION NUMBER: 60/277,337  
; PRIOR FILING DATE: 2001-03-20  
; NUMBER OF SEQ ID NOS: 159  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 86  
; LENGTH: 1531  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-939-853A-86

Query Match	79.4%	Score 7860;	DB 11;	Length 1531;
Best Local Similarity	100.0%	Pred. No. 0;	Mismatches 0;	Indels 0;
Matches 1531;	Conservative	0;		
Qy	1	MALRGFCSADGSDPLMDWNVTWNTSNPDTKCFQNTVLVWVPCFYLWVACFPFYLYLSRH	60	
Db	1	MALRGFCSADGSDPLMDWNVTWNTSNPDTKCFQNTVLVWVPCFYLWVACFPFYLYLSRH	60	
Qy	61	DRGYQMTPLNKTATLGLFWLWVCWADLFVSFWSRSGIFLAPVFLVSPITLLGLTTLLA	120	

Db 61 DRGYIQMTPLNKTALGFLMLIVCWADLFYSPWRSRGIFLAPVFLVSPPTLLGTTLLA 120  
 Qy 121 TELIQLERKGVSSGIMLTFWLVALCALAILRSKIMTALKEDAQVDLFRDITFYVYFS 180  
 Db 121 TELIQLERKGVSSGIMLTFWLVALCALAILRSKIMTALKEDAQVDLFRDITFYVYFS 180  
 Qy 181 LLLIQLVLSGSDRSPFLSETIHDNPPCPSSASFLSRITFWMTGLIVRGVQPLEGSD 240  
 Db 181 LLLIQLVLSGSDRSPFLSETIHDNPPCPSSASFLSRITFWMTGLIVRGVQPLEGSD 240  
 Qy 241 LMSLNKEDTSEQVVPVVLVKNWKKCAKTRKQPKVYVSSKDPKPKSSKVDANEVEAL 300  
 Db 241 LMSLNKEDTSEQVVPVVLVKNWKKCAKTRKQPKVYVSSKDPKPKSSKVDANEVEAL 300  
 Qy 301 IVKSPQKWNPSLFLVLYKTFPGPYFLMSPFFKAIHDLMMFSGPQILKLIKEVNDTKAPD 360  
 Db 301 IVKSPQKWNPSLFLVLYKTFPGPYFLMSPFFKAIHDLMMFSGPQILKLIKEVNDTKAPD 360  
 Qy 361 WQGYFYTVLLFTVACLTQLVLHQQYFHCIVSGMRKTAIVGAVYRKALVITNSARKSSTV 420  
 Db 361 WQGYFYTVLLFTVACLTQLVLHQQYFHCIVSGMRKTAIVGAVYRKALVITNSARKSSTV 420  
 Qy 421 GEIVNLMVDAQRFMDLTIYINMISAPLOVILALYLLNLPVSLAGVAVMLMVPVN 480  
 Db 421 GEIVNLMVDAQRFMDLTIYINMISAPLOVILALYLLNLPVSLAGVAVMLMVPVN 480  
 Qy 481 AVAMKTKTYQVAHMKSKDNRIKLMEILINGIKVLKVAWELAFKDKVLAIROBELKVLK 540  
 Db 481 AVAMKTKTYQVAHMKSKDNRIKLMEILINGIKVLKVAWELAFKDKVLAIROBELKVLK 540  
 Qy 541 KSAYLSAVGTFTWVCTPFLVALCTFAVYVTDENNILDAQTAFAVSLALFNILRFPNLILP 600  
 Db 541 KSAYLSAVGTFTWVCTPFLVALCTFAVYVTDENNILDAQTAFAVSLALFNILRFPNLILP 600  
 Qy 601 MYISSIVQASVLSKLRFLSHEELEPDSIERRPVKGGTNSITVRNATFTWARSDDPT 660  
 Db 601 MYISSIVQASVLSKLRFLSHEELEPDSIERRPVKGGTNSITVRNATFTWARSDDPT 660  
 Qy 661 LINGITFISPEGALVAVGVQGGKSSLLSALLAEMDKVEGHVIAIKGSVAYVPOQAWIOND 720  
 Db 661 LINGITFISPEGALVAVGVQGGKSSLLSALLAEMDKVEGHVIAIKGSVAYVPOQAWIOND 720  
 Qy 721 SIRENIFGCOLLEPYRVSIOACALLPDLPLSGDRTEIGEKGWNLSSGQKQVSLAR 780  
 Db 721 SIRENIFGCOLLEPYRVSIOACALLPDLPLSGDRTEIGEKGWNLSSGQKQVSLAR 780  
 Qy 781 AVYSNADILFDDPLSADVAHVCKHIFENVIGPKGMLKNKTRILVTHSMSYLPQVDVTV 840  
 Db 781 AVYSNADILFDDPLSADVAHVCKHIFENVIGPKGMLKNKTRILVTHSMSYLPQVDVTV 840  
 Qy 841 MSGGKISEMGSYQELLARDGAFELRTVYASTEQDAAENGVTGSGPGKEAKQWENG 900  
 Db 841 MSGGKISEMGSYQELLARDGAFELRTVYASTEQDAAENGVTGSGPGKEAKQWENG 900  
 Qy 901 LVTDAGKQLOQLSSSSSYSGDISRHNSHNSAELOKAEAKKEETWKLMEADKAQTVQVYL 960  
 Db 901 LVTDAGKQLOQLSSSSSYSGDISRHNSHNSAELOKAEAKKEETWKLMEADKAQTVQVYL 960  
 Qy 961 SVYDYMKAIGLIFSLFSLIFLPMCNHVSALASNYWLSLWTDPTVNGQEHKTVRLSYG 1020  
 Db 961 SVYDYMKAIGLIFSLFSLIFLPMCNHVSALASNYWLSLWTDPTVNGQEHKTVRLSYG 1020  
 Qy 1021 ALGISQGIAGVGSMAVSIIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFSEL 1080  
 Db 1021 ALGISQGIAGVGSMAVSIIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFSEL 1080  
 Qy 1081 DTVDNMIPEVIMFMGSLFNVIACIVILLATPAAIIIPPLGIYFPVQFYVASSRQL 1140  
 Db 1081 DTVDNMIPEVIMFMGSLFNVIACIVILLATPAAIIIPPLGIYFPVQFYVASSRQL 1140  
 Qy 1141 KRLESVSRSPVYSHENETLLGVSVIRAFEEQERFTHOSDLKVDENOKAYYSIVANRWLA 1200  
 Db 1141 KRLESVSRSPVYSHENETLLGVSVIRAFEEQERFTHOSDLKVDENOKAYYSIVANRWLA 1200

RESULT 2

US-10-667-891-6  
 ; Sequence 6, Application US/10667891  
 ; Publication No. US20040171024A1  
 ; GENERAL INFORMATION:

; APPLICANT: ROTH, CHARLES W.  
 ; APPLICANT: BREY, PAUL T.  
 ; APPLICANT: HOLM, INGE  
 ; APPLICANT: GRAILLES, MARINE  
 ; APPLICANT: RZHEITSKY, ANDREY  
 ; TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS IN DROSOPHILA AND  
 ; TITLE OF INVENTION: ANOPHELES  
 ; FILE REFERENCE: 03495.0294-00000  
 ; CURRENT APPLICATION NUMBER: US/10/667,891  
 ; PRIOR FILING DATE: 2003-09-23  
 ; PRIOR APPLICATION NUMBER: 60/413,469  
 ; PRIOR FILING DATE: 2002-09-26  
 ; NUMBER OF SEQ ID NOS: 76  
 ; SOFTWARE: Patent in Ver. 3.2  
 ; SEQ ID NO 6  
 ; LENGTH: 1531  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-667-891-6

Query Match 79.4%; Score 7860; DB 16; Length 1531;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MALRGFCSADGSDPLWDNVTWNTSNDFTKCFONTVLVWVPCFYLWACPPFFLYLSRH 60  
 Db 1 MALRGFCSADGSDPLWDNVTWNTSNDFTKCFONTVLVWVPCFYLWACPPFFLYLSRH 60  
 Qy 61 DRGYIQMTPLNKTALGFLMLIVCWADLFYSPWRSRGIFLAPVFLVSPPTLLGTTLLA 120  
 Db 61 DRGYIQMTPLNKTALGFLMLIVCWADLFYSPWRSRGIFLAPVFLVSPPTLLGTTLLA 120  
 Qy 121 TELIQLERKGVSSGIMLTFWLVALCALAILRSKIMTALKEDAQVDLFRDITFYVYFS 180  
 Db 121 TELIQLERKGVSSGIMLTFWLVALCALAILRSKIMTALKEDAQVDLFRDITFYVYFS 180  
 Qy 181 LLLIQLVLSGSDRSPFLSETIHDNPPCPSSASFLSRITFWMTGLIVRGVQPLEGSD 240  
 Db 181 LLLIQLVLSGSDRSPFLSETIHDNPPCPSSASFLSRITFWMTGLIVRGVQPLEGSD 240  
 Qy 241 LMSLNKEDTSEQVVPVVLVKNWKKCAKTRKQPKVYVSSKDPKPKSSKVDANEVEAL 300

Db 241 LWSLNKEDTSEQVFLVKNWKKCAKTRKQPKVYSSKDPAPQPKSSKVDANEVEAL 300  
 Qy 301 IVKSPQKWNPSLFLVLYKTFPGPYFLMGFFPKAIHDLMMFSGPOLKLLIKFVNDTKAPD 360  
 Db 301 IVKSPQKWNPSLFLVLYKTFPGPYFLMGFFPKAIHDLMMFSGPOLKLLIKFVNDTKAPD 360  
 Qy 361 WQGYFYTVLLFVTTACLOTLVHLHQYFHCIVGSGMRKTAIVGAVYRKALVITNSARKSSTV 420  
 Db 361 WQGYFYTVLLFVTTACLOTLVHLHQYFHCIVGSGMRKTAIVGAVYRKALVITNSARKSSTV 420  
 Qy 421 GEIVNLMSDAQRFDMLATYINMIWSAPLOVILALYLLWNLGSPSLAGVAVMLMVPVN 480  
 Db 421 GEIVNLMSDAQRFDMLATYINMIWSAPLOVILALYLLWNLGSPSLAGVAVMLMVPVN 480  
 Qy 481 AVAMKTKTYQVAHMKSKDNRIKLWNEILNGIKVLKYLWELAFKDKVLAIROBELKVLK 540  
 Db 481 AVAMKTKTYQVAHMKSKDNRIKLWNEILNGIKVLKYLWELAFKDKVLAIROBELKVLK 540  
 Qy 541 KSAYLSAVGTFTWCTPPLVALCTFAVYVTTDENNILDQAQTAFAVSLALFNILRPFNLILP 600  
 Db 541 KSAYLSAVGTFTWCTPPLVALCTFAVYVTTDENNILDQAQTAFAVSLALFNILRPFNLILP 600  
 Qy 601 MVISSIVQASVLSKRLRIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWARSDDPT 660  
 Db 601 MVISSIVQASVLSKRLRIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWARSDDPT 660  
 Qy 661 LNGITFSP1PEGALVAVGVGQCKSLLSALLAENDKVEGHVAIKGSVAVYVPOQAWIQND 720  
 Db 661 LNGITFSP1PEGALVAVGVGQCKSLLSALLAENDKVEGHVAIKGSVAVYVPOQAWIQND 720  
 Qy 721 SURENIFGQCLEEPYRSVIOACALLPDLPLSGDRTEIGEKNVLSGGQKQVSLAR 780  
 Db 721 SURENIFGQCLEEPYRSVIOACALLPDLPLSGDRTEIGEKNVLSGGQKQVSLAR 780  
 Qy 781 AVYSNADIYLPDPLSADVAHVGHKHFENVIGPKGMLKNKTRILVTHSMSYLPQVDVLIIV 840  
 Db 781 AVYSNADIYLPDPLSADVAHVGHKHFENVIGPKGMLKNKTRILVTHSMSYLPQVDVLIIV 840  
 Qy 841 MSGGKISEMGYSQELLARDGAFELRTVYASTEQODAEENGTVGSGPKAKOMENGM 900  
 Db 841 MSGGKISEMGYSQELLARDGAFELRTVYASTEQODAEENGTVGSGPKAKOMENGM 900  
 Qy 901 LVTDSAGKQLOQLSSSSSYSGDISRHHNSTAELQKAEAKKEBTWKLMEADKAQGTQVKL 960  
 Db 901 LVTDSAGKQLOQLSSSSSYSGDISRHHNSTAELQKAEAKKEBTWKLMEADKAQGTQVKL 960  
 Qy 961 SYTDYMKAIGLFISFLFPMCHVSNALSNYWSLWTDPIVNGTQEBTKVRLSVYG 1020  
 Db 961 SYTDYMKAIGLFISFLFPMCHVSNALSNYWSLWTDPIVNGTQEBTKVRLSVYG 1020  
 Qy 1021 ALGISQGIATVFCYSMAVSTGGILASRCLHVDLHSLRSPMSFFERTPSGNLVNRFSEL 1080  
 Db 1021 ALGISQGIATVFCYSMAVSTGGILASRCLHVDLHSLRSPMSFFERTPSGNLVNRFSEL 1080  
 Qy 1081 DTVDMSIPEVIVKFMGSLFNVICACIVILLATPIAAIIPPLGLIYFFVQRYVASSRQL 1140  
 Db 1081 DTVDMSIPEVIVKFMGSLFNVICACIVILLATPIAAIIPPLGLIYFFVQRYVASSRQL 1140  
 Qy 1141 KRLESVSRSPVYSHNETLLGVSVTRAFEEQBRFTHQSDLKVDENQKAYPYSIVANRWLA 1200  
 Db 1141 KRLESVSRSPVYSHNETLLGVSVTRAFEEQBRFTHQSDLKVDENQKAYPYSIVANRWLA 1200  
 Qy 1201 VLECEVGNICVILFAALFAVISHRSUSAGLVGVSYSQVTTYLNLWRMSSEMETNIVA 1260  
 Db 1201 VLECEVGNICVILFAALFAVISHRSUSAGLVGVSYSQVTTYLNLWRMSSEMETNIVA 1260  
 Qy 1261 VERLKEYSETEKAPWQIOETAPPSSWPQGVREFRNYCLRYREDLDFVLRHINTVINGG 1320  
 Db 1261 VERLKEYSETEKAPWQIOETAPPSSWPQGVREFRNYCLRYREDLDFVLRHINTVINGG 1320  
 Qy 1321 EXKVGIVGRTGAGKSSLTGLFRINSAEGEIIIDGINIAKIGHDLRFRKTIIPQDPVLF 1380  
 Db 1321 EXKVGIVGRTGAGKSSLTGLFRINSAEGEIIIDGINIAKIGHDLRFRKTIIPQDPVLF 1380

Qy 1381 SGLRMLNDPFSQYSDDEVWTSLELAHLKDFVSALPDKLDHECABEGGENLSVGORQLVCL 1440  
 Db 1381 SGLRMLNDPFSQYSDDEVWTSLELAHLKDFVSALPDKLDHECABEGGENLSVGORQLVCL 1440  
 Qy 1441 ARALLRKTIIULVDEBATAAVIDETDDLLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 Db 1441 ARALLRKTIIULVDEBATAAVIDETDDLLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 Qy 1501 DKGEIQEYGAQPSDILLQORGLFYSMAKDAGLV 1531  
 Db 1501 DKGEIQEYGAQPSDILLQORGLFYSMAKDAGLV 1531

RESULT 3

US-10-484-577-678  
 ; Sequence 678, Application US/10484577  
 ; Publication No. US20050032724A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: EPIDAUROS Biotechnologie Aktiengesellschaft  
 ; TITLE OF INVENTION: Means and methods for improved treatment of cancer based on Ugr1  
 ; FILE REFERENCE: F2285PCT-1  
 ; CURRENT APPLICATION NUMBER: US/10/484,577  
 ; CURRENT FILING DATE: 2004-01-22  
 ; PRIOR APPLICATION NUMBER: PCT/EP 02/08220  
 ; PRIOR FILING DATE: 2002-07-23  
 ; PRIOR APPLICATION NUMBER: EP 01 11 7608.8  
 ; PRIOR FILING DATE: 2001-07-23  
 ; PRIOR APPLICATION NUMBER: EP 02011710.7  
 ; PRIOR FILING DATE: 2002-05-24  
 ; NUMBER OF SEQ ID NOS: 683  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 678  
 ; LENGTH: 1531  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-484-577-678

Query Match 79.4%; Score 7860; DB 17; Length 1531;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MALRGFCSDGSDPLMDNVTWNTSNPDFTKCFQNTVLVWVPCFYLMACFPFYLYLSRH 60  
 Db 1 MALRGFCSDGSDPLMDNVTWNTSNPDFTKCFQNTVLVWVPCFYLMACFPFYLYLSRH 60  
 Qy 61 DRGIQMTPIKTKTALGFLWVWADLFSFWERSRGIPLAPVFLVSPTLIGITTLIA 120  
 Db 61 DRGIQMTPIKTKTALGFLWVWADLFSFWERSRGIPLAPVFLVSPTLIGITTLIA 120  
 Qy 121 TFLIQLERRKGVQSSGIMLTFWLVALVCAALILRSKIMTALKEDAQVDLFRDITFYVYPS 180  
 Db 121 TFLIQLERRKGVQSSGIMLTFWLVALVCAALILRSKIMTALKEDAQVDLFRDITFYVYPS 180  
 Qy 181 LLLIQLVLSGSDRSPSFSETHDPNCPSSASFLSRITFWITGLIVRGVQPLEGSD 240  
 Db 181 LLLIQLVLSGSDRSPSFSETHDPNCPSSASFLSRITFWITGLIVRGVQPLEGSD 240  
 Qy 241 LWSLNKEDTSQVVPVLVKNWKKCAKTRKQPKVYSSKDPAPQPKSSKVDANEVEAL 300  
 Db 241 LWSLNKEDTSQVVPVLVKNWKKCAKTRKQPKVYSSKDPAPQPKSSKVDANEVEAL 300  
 Qy 301 IVKSPQKWNPSLFLVLYKTFPGPYFLMGFFPKAIHDLMMFSGPOLKLLIKFVNDTKAPD 360  
 Db 301 IVKSPQKWNPSLFLVLYKTFPGPYFLMGFFPKAIHDLMMFSGPOLKLLIKFVNDTKAPD 360  
 Qy 361 WQGYFYTVLLFVTTACLOTLVHLHQYFHCIVGSGMRKTAIVGAVYRKALVITNSARKSSTV 420  
 Db 361 WQGYFYTVLLFVTTACLOTLVHLHQYFHCIVGSGMRKTAIVGAVYRKALVITNSARKSSTV 420  
 Qy 421 GEIVNLMSDAQRFDMLATYINMIWSAPLOVILALYLLWNLGSPSLAGVAVMLMVPVN 480  
 Db 421 GEIVNLMSDAQRFDMLATYINMIWSAPLOVILALYLLWNLGSPSLAGVAVMLMVPVN 480

481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAWELAFKDKVLAIROBELKVLK 540  
 Db AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAWELAFKDKVLAIROBELKVLK 540  
 541 KSAYLSAVGTFWTCVTPFLVALCTFAVYVTTIDENNILDAQTAQFVSLALFNILRFFLNILP 600  
 Db KSAYLSAVGTFWTCVTPFLVALCTFAVYVTTIDENNILDAQTAQFVSLALFNILRFFLNILP 600  
 601 MWISSIVQASVSLKRLRIFLSHEELEPDSDIERRPVKDGGTNSITVRNATFTWASDPT 660  
 Db MWISSIVQASVSLKRLRIFLSHEELEPDSDIERRPVKDGGTNSITVRNATFTWASDPT 660  
 661 LNGITFISPEGALVAVVGQVGGKSLLSALLAEMDKVEGHVAIKGSVAYVPOQAWIQND 720  
 Db LNGITFISPEGALVAVVGQVGGKSLLSALLAEMDKVEGHVAIKGSVAYVPOQAWIQND 720  
 721 SURENIFGCOLLEPYRYSVIOACALLPDLTEILPSGDRTEIGEGKGNLSGGQKQVSLAR 780  
 Db SURENIFGCOLLEPYRYSVIOACALLPDLTEILPSGDRTEIGEGKGNLSGGQKQVSLAR 780  
 781 AVYSNADIYLPDDPLSADVAHVGHKIFENVIGPKGMLKNKTRILVTHSMSYLPQVDVIV 840  
 Db AVYSNADIYLPDDPLSADVAHVGHKIFENVIGPKGMLKNKTRILVTHSMSYLPQVDVIV 840  
 841 MSGGKISEMGYSQELLARDGAFELRTYASTEQDQAEENGVTGVSQKQKQWENG 900  
 Db MSGGKISEMGYSQELLARDGAFELRTYASTEQDQAEENGVTGVSQKQKQWENG 900  
 901 LVTDAGKOLQOLSSSSSYSGDISRHNSSTAELQAKAEKKEETWKLMEADKAQGVKL 960  
 Db LVTDAGKOLQOLSSSSSYSGDISRHNSSTAELQAKAEKKEETWKLMEADKAQGVKL 960  
 961 SVYDYMKAIGLFIISFLIFPMCHVNSALASNYWLSLWTDPIVNGQEHKVRSLSYG 1020  
 Db SVYDYMKAIGLFIISFLIFPMCHVNSALASNYWLSLWTDPIVNGQEHKVRSLSYG 1020  
 1021 ALGISQGLAVFGYSMAVSIIGTILASRLVLDLHLSILSPMSFFERTSGNIVNRFKSEL 1080  
 Db ALGISQGLAVFGYSMAVSIIGTILASRLVLDLHLSILSPMSFFERTSGNIVNRFKSEL 1080  
 1081 DTVDMSIPEVIKMFNGSLFNVIGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQL 1140  
 Db DTVDMSIPEVIKMFNGSLFNVIGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQL 1140  
 1141 KELESVSRSPVYSHENETLLGVSVIRAEEOERFIHQSDLKVDENOKAYYPSIVANRWLA 1200  
 Db KELESVSRSPVYSHENETLLGVSVIRAEEOERFIHQSDLKVDENOKAYYPSIVANRWLA 1200  
 1201 VRLECVGNICVILFAALFAVIRHSLSAGLVGLSVSYSLQVTTYLNWLVRMSSEMETNIVA 1260  
 Db VRLECVGNICVILFAALFAVIRHSLSAGLVGLSVSYSLQVTTYLNWLVRMSSEMETNIVA 1260  
 1261 VERLKEYSETEKAPWQIOETAPSSWPQVGRVEFRNYCLRYREDLDLFLVRHINVTINGG 1320  
 Db VERLKEYSETEKAPWQIOETAPSSWPQVGRVEFRNYCLRYREDLDLFLVRHINVTINGG 1320  
 1321 EKVGIUGRTGAGKSLTGLFRINESAGEEIIIDGINIAKTLGHLDRKFIITIIIPQDPVLF 1380  
 Db EKVGIUGRTGAGKSLTGLFRINESAGEEIIIDGINIAKTLGHLDRKFIITIIIPQDPVLF 1380  
 1381 SGSLRMNLDPPFSQYSDDEEVTSLAHLKDFVSALPDKLDHECAGGNGENLVGQRLVCL 1440  
 Db SGSLRMNLDPPFSQYSDDEEVTSLAHLKDFVSALPDKLDHECAGGNGENLVGQRLVCL 1440  
 1441 ARALLRKTILVLDEATAVLETDLLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRIVL 1500  
 Db ARALLRKTILVLDEATAVLETDLLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRIVL 1500  
 1501 DKGEIQEYGAUSDLLQKGLFYSAKADAGLV 1531  
 Db DKGEIQEYGAUSDLLQKGLFYSAKADAGLV 1531

RESULT 4  
 US-09-939-853A-87  
 ; Sequence 87, Application US/09939853A  
 ; Publication No. US20040039163A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Burgess et al.  
 ; TITLE OF INVENTION: No. US20040039163A1el Proteins and Nucleic Acids Encoding Same  
 ; FILE REFERENCE: 21402-099  
 ; CURRENT APPLICATION NUMBER: US/09/939,853A  
 ; CURRENT FILING DATE: 2001-08-27  
 ; PRIOR APPLICATION NUMBER: 60/228,191  
 ; PRIOR FILING DATE: 2000-08-25  
 ; PRIOR APPLICATION NUMBER: 60/267,300  
 ; PRIOR FILING DATE: 2001-02-08  
 ; PRIOR APPLICATION NUMBER: 60/269,961  
 ; PRIOR FILING DATE: 2001-02-20  
 ; PRIOR APPLICATION NUMBER: 60/277,337  
 ; PRIOR FILING DATE: 2001-03-20  
 ; NUMBER OF SEQ ID NOS: 159  
 ; SOFTWARE: Patent in Ver. 2.1  
 ; SEQ ID NO 87  
 ; LENGTH: 1515  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-939-853A-87

Query Match 78.5%; Score 7769; DB 11; Length 1515;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1515; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 DWNTWNTSNPDPFKCFQNTVLVWVPCFYLWACFPFYLILSRDRGVQMTPLNKTKTA 76  
 Db 1 DWNTWNTSNPDPFKCFQNTVLVWVPCFYLWACFPFYLILSRDRGVQMTPLNKTKTA 60

Qy 77 LGFLLLWVCVADLWLFYFWERSRGIFLAPVFLVSTLIGITTLATLFIQLERRRGVQSSG 136  
 Db 61 LGFLLLWVCVADLWLFYFWERSRGIFLAPVFLVSTLIGITTLATLFIQLERRRGVQSSG 120

Qy 137 IMLFTWLVALCALAILRSKIMTALKEDAQVDLFRDITFYVYFSLLIQLVLSCFSDRSP 196  
 Db 121 IMLFTWLVALCALAILRSKIMTALKEDAQVDLFRDITFYVYFSLLIQLVLSCFSDRSP 180

Qy 197 LFSSTIHDNPPCPSSASFLSRITFWITGLIVRGYQPLEGSLWSINKEDTSEOVVVP 256  
 Db 181 LFSSTIHDNPPCPSSASFLSRITFWITGLIVRGYQPLEGSLWSINKEDTSEOVVVP 240

Qy 257 LVKNWKECAKTRKQPVKVYVSSKDPAPKSESSKVDANEEVEALIVKSPQKWNPSLFKV 316  
 Db 241 LVKNWKECAKTRKQPVKVYVSSKDPAPKSESSKVDANEEVEALIVKSPQKWNPSLFKV 300

Qy 317 LYKTFGYPFLMSPFFKAIHDMFSGPQILKLLIKFVNDTKAPDWQGYFYTVLLFVTAQL 376  
 Db 301 LYKTFGYPFLMSPFFKAIHDMFSGPQILKLLIKFVNDTKAPDWQGYFYTVLLFVTAQL 360

Qy 377 QTLVLHQYFHCFSVGNRIKTAVIGAVYRKALVTNSARKSSTVGEIVNLSMVDQAQRPMD 436  
 Db 361 QTLVLHQYFHCFSVGNRIKTAVIGAVYRKALVTNSARKSSTVGEIVNLSMVDQAQRPMD 420

Qy 437 LATYINNIWSAPLOVILALYLLMLNLGSPVLAVVAVMLVMPVNAVMAKTKTYQVAHMK 496  
 Db 421 LATYINNIWSAPLOVILALYLLMLNLGSPVLAVVAVMLVMPVNAVMAKTKTYQVAHMK 480

Qy 497 SKNRIKLMNEILNGIKVLKYAWELAFKDKVLAIROBELKVLKKSAYLSAVGFTTWCT 556  
 Db 481 SKNRIKLMNEILNGIKVLKYAWELAFKDKVLAIROBELKVLKKSAYLSAVGFTTWCT 540

Qy 557 PFLVALCTFAVYVTTIDENNILDAQTAQFVSLALFNILRFFLNILPMTWISSIVQASVSLKRL 616  
 Db 541 PFLVALCTFAVYVTTIDENNILDAQTAQFVSLALFNILRFFLNILPMTWISSIVQASVSLKRL 600

Qy 617 RIFLSHEELEPDSDIERRPVKDGGTNSITVRNATFTWASDPTLNGITTSIPGALVAV 676  
 Db 617 RIFLSHEELEPDSDIERRPVKDGGTNSITVRNATFTWASDPTLNGITTSIPGALVAV 676

```

Db 601 RIFLSHELEPDSIERPVDGGTNSITVRNATFTWAKSDPPTLNGITFTSIPGALVAV 660
QY 677 VQVGGCKSSLLSALLAEMDKVEGHVAIKGSVAYVPOQAWIQNDLSRENILPGCQLEBPY 736
Db 661 VQVGGCKSSLLSALLAEMDKVEGHVAIKGSVAYVPOQAWIQNDLSRENILPGCQLEBPY 720
QY 737 YRSVTOACALLPDLLEILPSGDRTEIGEGVNLGGQKQKRVSLARAVYNSADIYLFDDPLS 796
Db 721 YRSVTOACALLPDLLEILPSGDRTEIGEGVNLGGQKQKRVSLARAVYNSADIYLFDDPLS 780
QY 797 AYDAHVGHIFENVIGPKMKNKTRILVTHSMVLPQVDVILVMSGGKISEMGSYQELL 856
Db 781 AYDAHVGHIFENVIGPKMKNKTRILVTHSMVLPQVDVILVMSGGKISEMGSYQELL 840
QY 857 ARDGAFAEFLRTYASTEQDQAEENGVTGSGPGKEAKQKMGMLVTDTSAGKQLOQLSS 916
Db 841 ARDGAFAEFLRTYASTEQDQAEENGVTGSGPGKEAKQKMGMLVTDTSAGKQLOQLSS 900
QY 917 SSSYSGDISRHHNSTAELQKABAKKEETWKLMEADKAOTGVKLSVYWDYMKAIGLFISF 976
Db 901 SSSYSGDISRHHNSTAELQKABAKKEETWKLMEADKAOTGVKLSVYWDYMKAIGLFISF 960
QY 977 LSIPLFCMCHVSALASNYWLSLWTDPIVNGTOEHTKVLRSVYGALGISOGIAVFGYSA 1036
Db 961 LSIPLFCMCHVSALASNYWLSLWTDPIVNGTOEHTKVLRSVYGALGISOGIAVFGYSA 1020
QY 1037 VSGIGILASRCHLVLDLHILSPMSFFERTPSGNLVNRFKSELDTVDOSMIPEVIMPMFG 1096
Db 1021 VSGIGILASRCHLVLDLHILSPMSFFERTPSGNLVNRFKSELDTVDOSMIPEVIMPMFG 1080
QY 1097 SLFNVIGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQLKRLSVSRSPVYSHFN 1156
Db 1081 SLFNVIGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQLKRLSVSRSPVYSHFN 1140
QY 1157 ETLGLSVIRAFEEQERFTHQSDLVKVDENOKAYPSIVANRWLAVRLECVGNCIVLFAAL 1216
Db 1141 ETLGLSVIRAFEEQERFTHQSDLVKVDENOKAYPSIVANRWLAVRLECVGNCIVLFAAL 1200
QY 1217 FAVISRHSLSAGLVLSYSYLSQVTTYLNWLRMSSEMETNIVAVERLKEYSETKEAPW 1276
Db 1201 FAVISRHSLSAGLVLSYSYLSQVTTYLNWLRMSSEMETNIVAVERLKEYSETKEAPW 1260
QY 1277 QIOTAPPSSWPQGVREFRNYCLARYREDLDFVLRHINVTYNGGKVGIVGRTGAGKSSL 1336
Db 1261 QIOTAPPSSWPQGVREFRNYCLARYREDLDFVLRHINVTYNGGKVGIVGRTGAGKSSL 1320
QY 1337 TLGLFRINESAGEIIIDGINIAKIGLHDLRPKITIIIPQDPVLFSGSLBMNLDPSQYSD 1396
Db 1321 TLGLFRINESAGEIIIDGINIAKIGLHDLRPKITIIIPQDPVLFSGSLBMNLDPSQYSD 1380
QY 1397 EEVWTSLELAHLKDFVSALPDKLDHECAGGENLSVGQRLVCLARALLRKTILVLDEA 1456
Db 1381 EEVWTSLELAHLKDFVSALPDKLDHECAGGENLSVGQRLVCLARALLRKTILVLDEA 1440
QY 1457 TAAVDLETDDLQISTIRTOFECTVLTIAHRLNTIMDYTRVIVLDKGETQYEGAPSDLLQ 1516
Db 1441 TAAVDLETDDLQISTIRTOFECTVLTIAHRLNTIMDYTRVIVLDKGETQYEGAPSDLLQ 1500
QY 1517 QRGLFYSMAKDAGLV 1531
Db 1501 QRGLFYSMAKDAGLV 1515

```

RESULT 5

```

; Sequence 1718, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.

```

```

; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1718
; LENGTH: 1388
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-408-765A-1718

Query Match 71.1%; Score 7037; DB 16; Length 1388;
Best Local Similarity 97.0%; Pred. No. 0;
Matches 1387; Conservative 0; Mismatches 1; Indels 42; Gaps 1;

QY 67 MTPLNKTKTALGFLLMIVCWADLFYSPWERSRGIFLAPVFLVSPPTLLGTTLLATFLIQL 126
Db 1 MTPLNKTKTALGFLLMIVCWADLFYSPWERSRGIFLAPVFLVSPPTLLGTTLLATFLIQL 60
QY 127 ERRKGVQSSGIMLTFWLVALVCALAILRSKIMTALKEDAQVDLPRDITFYVYFSLLLQL 186
Db 61 ERRKGVQSSGIMLTFWLVALVCALAILRSKIMTALKE ----- 97
QY 187 VLSCFSDRSPLFSETHDPNCPSSASFLSRIITFWMITGLIVRGYROPLEGSDLSLNK 246
Db 98 -----NPCSSASFLSRIITFWMITGLIVRGYROPLEGSDLSLNK 138
QY 247 EDTSQVVPVLVKNWKCECAKTRQPVKVYSSKDPAPQKSSKYDANEEVEALIVKSPQ 306
Db 139 EDTSQVVPVLVKNWKCECAKTRQPVKVYSSKDPAPQKSSKYDANEEVEALIVKSPQ 198
QY 307 KEWNPSPKVLKTPGPFPLMSFFPKAIDHLMFSGPQILKLIKFVNDTKAPDMQGYFY 366
Db 199 KEWNPSPKVLKTPGPFPLMSFFPKAIDHLMFSGPQILKLIKFVNDTKAPDMQGYFY 258
QY 367 TVLLFVTTACLOTVLVHGFHICFVSGMRIKTAVICAVYRKALVITNSARKSSTVGEIVNL 426
Db 259 TVLLFVTTACLOTVLVHGFHICFVSGMRIKTAVICAVYRKALVITNSARKSSTVGEIVNL 318
QY 427 MSVDAQRFMDLATYINMIWSAPLOVILALYLLWNLGSPVLGAVAVMVLMPVNAVMAK 486
Db 319 MSVDAQRFMDLATYINMIWSAPLOVILALYLLWNLGSPVLGAVAVMVLMPVNAVMAK 378
QY 487 TKTQVAHMKSDNRIKLMNBIINGIKVLKYAWELAFKDKVLAIROBELKVLKKSAYLS 546
Db 379 TKTQVAHMKSDNRIKLMNBIINGIKVLKYAWELAFKDKVLAIROBELKVLKKSAYLS 438
QY 547 AVGTFTWCTPFLVALCTFAVVYTTIDENNILDAQTAFLSLALFNILRPPNLPVVISSI 606
Db 439 AVGTFTWCTPFLVALCTFAVVYTTIDENNILDAQTAFLSLALFNILRPPNLPVVISSI 498
QY 607 VQASVSLKRLRIFLSHEELEPDSIERPVDGGTNSITVRNATFTWAKSDPPTLNGITFT 666
Db 499 VQASVSLKRLRIFLSHEELEPDSIERPVDGGTNSITVRNATFTWAKSDPPTLNGITFT 558
QY 667 SIPEGALVAVVQVCGCKSSLLSALLAEMDKVEGHVAIKGSVAYVPOQAWIQNDLSREN 726
Db 559 SIPEGALVAVVQVCGCKSSLLSALLAEMDKVEGHVAIKGSVAYVPOQAWIQNDLSREN 618
QY 727 LFGCOLLEBPYRSVTOACALLPDLLEILPSGDRTEIGEGVNLGGQKQKRVSLARAVYSA 786
Db 619 LFGCOLLEBPYRSVTOACALLPDLLEILPSGDRTEIGEGVNLGGQKQKRVSLARAVYSA 678
QY 787 DIYLFDDPLSADVAHVGHIFENVIGPKMKNKTRILVTHSMVLPQVDVILVMSGGKI 846
Db 679 DIYLFDDPLSADVAHVGHIFENVIGPKMKNKTRILVTHSMVLPQVDVILVMSGGKI 738
QY 847 SEMGSYQELLARDGAFABFLRTYASTEQDQAEENGVTGSGPGKEAKQKMGMLVTD 906

```

Db 739 SEMGYSQELLARDGAFABFLRTYASTEQEADAEENGVTGSGFGKEAKQMGMLVTD5A 798  
 Qy 907 GKOLQRLSSSSSYSGDISRHNSHTAELOKAKKEETWKLMEADKAQOGVQLSVYWDY 966  
 Db 799 GKOLQRLSSSSSYSGDISRHNSHTAELOKAKKEETWKLMEADKAQOGVQLSVYWDY 858  
 Qy 967 MKAGLGFISFLSFLFMCNHNVSALASNYWLSLWTDPIVNGTQETHKVRSLVYGALGISQ 1026  
 Db 859 MKAGLGFISFLSFLFMCNHNVSALASNYWLSLWTDPIVNGTQETHKVRSLVYGALGISQ 918  
 Qy 1027 GIAVFGYSMAVIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVRFSKELDTVDSM 1086  
 Db 919 GIAVFGYSMAVIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVRFSKELDTVDSM 978  
 Qy 1087 IPEVIMKMGSLFNVTGACIVILLATPTAAIIIPGLIYFFVQRFYVASSRQLKRLSV 1146  
 Db 979 IPEVIMKMGSLFNVTGACIVILLATPTAAIIIPGLIYFFVQRFYVASSRQLKRLSV 1038  
 Qy 1147 SRSPVYSHNETLLGVSVIRAFEEQERFTHQSDLVKVDENOKAYPSIVANRWLAVRLECV 1206  
 Db 1039 SRSPVYSHNETLLGVSVIRAFEEQERFTHQSDLVKVDENOKAYPSIVANRWLAVRLECV 1098  
 Qy 1207 GNCIVLFAALFAVIRSHLSAGLVGSYSISQVTTYLNWLVRMSSEMETNIVAVERLKE 1266  
 Db 1099 GNCIVLFAALFAVIRSHLSAGLVGSYSISQVTTYLNWLVRMSSEMETNIVAVERLKE 1158  
 Qy 1267 YSETEKEAPWQIQETAPSSWPQGRVFRNVCYLYREDLDPVLRHINVTINGGKGVIV 1326  
 Db 1159 YSETEKEAPWQIQETAPSSWPQGRVFRNVCYLYREDLDPVLRHINVTINGGKGVIV 1218  
 Qy 1327 GRTGAGKSSLTGLFRINESAGEIIDIINIAGIHLDRPKITIIIPQDPVLFGSLRM 1386  
 Db 1219 GRTGAGKSSLTGLFRINESAGEIIDIINIAGIHLDRPKITIIIPQDPVLFGSLRM 1278  
 Qy 1387 NLDPSQYSDREWVTSLELAHLKDFVSALPKLDHECAEGGENLSVQORQVLCIARALLR 1446  
 Db 1279 NLDPSQYSDREWVTSLELAHLKDFVSALPKLDHECAEGGENLSVQORQVLCIARALLR 1338  
 Qy 1447 KTKILVLDEATAVLETDLLIQSTIRIQFEDCTVLTIAHRLNTIMDYTR 1496  
 Db 1339 KTKILVLDEATAVLETDLLIQSTIRIQFEDCTVLTIAHRLNTIMDYTR 1388

RESULT 6  
 US-09-939-853A-88  
 ; Sequence 88, Application US/09939853A  
 ; Publication No. US20040039163A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Burgess et al.  
 ; TITLE OF INVENTION: No. US20040039163A1el Proteins and Nucleic Acids Encoding Same  
 ; FILE REFERENCE: 21402-099  
 ; CURRENT APPLICATION NUMBER: US/09/939,853A  
 ; PRIORITY FILING DATE: 2001-08-27  
 ; PRIOR APPLICATION NUMBER: 60/228,191  
 ; PRIOR FILING DATE: 2000-08-25  
 ; PRIOR APPLICATION NUMBER: 60/267,300  
 ; PRIOR FILING DATE: 2001-02-08  
 ; PRIOR APPLICATION NUMBER: 60/269,961  
 ; PRIOR FILING DATE: 2001-02-20  
 ; PRIOR APPLICATION NUMBER: 60/277,337  
 ; PRIOR FILING DATE: 2001-03-20  
 ; NUMBER OF SEQ ID NOS: 159  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 88  
 ; LENGTH: 1528  
 ; TYPE: PRT  
 ; ORGANISM: Mus musculus  
 US-09-939-853A-88

Query Match 70.7%; Score 7002.5; DB 11; Length 1528;  
 Best Local Similarity 88.0%; Pred. No. 0;  
 Matches 1349; Conservative 102; Mismatches 75; Indels 7; Gaps 5;

Qy 1 MALRGFCSDGSDPLWDNNTVNTSNPDFTKCFONTVLVWVPCFYLWACFPFFFLYLSRH 60  
 Db 1 MAURSFCSADGSDPLWDNNTVNTSNPDFTKCFONTVLVWVPCFYLWACFPFFFLYLSRH 60  
 Qy 61 DRGVIQMTPLNKTKTALGFLWIVCWADLFYSFWERSRGIFLAPVFLVPSYLLGTTLLA 120  
 Db 61 DRGVIQMTPLNKTKTALGFLWIVCWADLFYSFWERSRGIFLAPVFLVPSYLLGTTLLA 120  
 Qy 121 TFLIQLERRKGVQSSGIMLTFWLVALCALAILRSKIMTALKEDAQVDLFRDITFYVYFS 180  
 Db 121 TFLIQLERRKGVQSSGIMLTFWLVALCALAILRSKIMTALKEDAQVDLFRDITFYVYFS 180  
 Qy 181 LLLIQLVLSCSFSDPLSETHDNPCESSASFLSRIETFWMITGLVIRGVRQPLESD 240  
 Db 181 LVLVOLVLSCSFSDPLSETHDNPCESSASFLSRIETFWMITGLVIRGVRQPLESD 240  
 Qy 241 LWSLNKEDTSQVVPVVLVKNWKECAKTRKQPVKVYVS-SKDPAPQPKESSKVDAEVEEA 299  
 Db 241 LWSLNKEDTSQVVPVVLVKNWKECAKTRKQPVKVYVS-SKDPAPQPKESSKVDAEVEEA 299  
 Qy 300 LIVSPQKEMNPSLKVLYKTFPGYFLMSFFFKAIHDLMPSPGPOILKLLIKFVNDTKAP 359  
 Db 301 LIVSPKHDBPSLFKVLYKTFPGYFLMSFLYKALHDLMPGPKILELIINFVNDREAP 360  
 Qy 360 DWQGVYTVLLFVTAQTLVHOYEHICFVSGMIRIKTAVIGAVYRKALVTNSAKSST 419  
 Db 361 DWQGVYTVLLFVTAQTLVHOYEHICFVSGMIRIKTAVIGAVYRKALVTNSAKSST 420  
 Qy 420 VGEIVNLSVDAQRFMDLATYINNIWSAPQVILALYLLMLNLGSPVLGAVVAVMLVVPV 479  
 Db 421 VGEIVNLSVDAQRFMDLATYINNIWSAPQVILALYLLMLNLGSPVLGAVVAVMLVVPV 480  
 Qy 480 NAVMAMTKTYQVAHMKSKONRIKLMNEILNGIKVLKYAWELAFKQVLAIROEELKVL 539  
 Db 481 NAVMAMTKTYQVAHMKSKONRIKLMNEILNGIKVLKYAWELAFKQVLAIROEELKVL 540  
 Qy 540 KKSAYLSAVGFTWVCTPFLVALCTFAVYVTDENNILDACTAFVSLALFNILFPNLL 599  
 Db 541 KKSAYLSAVGFTWVCTPFLVALCTFAVYVTDENNILDACTAFVSLALFNILFPNLL 600  
 Qy 600 PMWISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKGGGTNSITVRNATFTTWARSDDP 659  
 Db 601 PMWISSIVQASVSLKRLRIFLSHEELEPDSIERRSIKSGEG-NSITVKNATFTTWARGEP 659  
 Qy 660 TLNGITFISIEGALVAVVGVQCGKSSILSALLAEMDKVEGHVAIKSVAVVPOQAIQ 719  
 Db 660 TLNGITFISIEGALVAVVGVQCGKSSILSALLAEMDKVEGHVAIKSVAVVPOQAIQ 719  
 Qy 720 DSLRENILFGQLEPEYRSVIOACALLPDLLEILPSGDRTEIGKGVNLSGGQKQVSLA 779  
 Db 720 DSLRENILFGHPLQENYKAVMEACALLPDLLEILPSGDRTEIGKGVNLSGGQKQVSLA 779  
 Qy 780 RAVYSNADIVLFDPLSAVDAHVGHKIFENVIGPKGMLKNKTRILVTHSHSYLPOVDVII 839  
 Db 780 RAVYSNADIVLFDPLSAVDAHVGHKIFENVIGPKGMLKNKTRILVTHSHSYLPOVDVII 839  
 Qy 840 VMGGKISEMGYSQELLARDGAFABFLRTYASTEQEADAEENGVTGSGFGKEAKQMGML 899  
 Db 840 VMGGKISEMGYSQELLARDGAFABFLRTYASTEQEADAEENGVTGSGFGKEAKQMGML 896  
 Qy 900 MLVTD5AGKQLQRLSSSSSYSGDISRHNSHTAELOKAKKEETWKLMEADKAQOGVQ 959  
 Db 897 MLVTD5AGKQLQRLSSSSSYSGDISRHNSHTAELOKAKKEETWKLMEADKAQOGVQ 955  
 Qy 960 LSVTWDYNKAIGLFISFLSFLFMCNHNVSALASNYWLSLWTDPIVNGTQETHKVRSLV 1018  
 Db 956 LSVTWDYNKAIGLFISFLSFLFMCNHNVSALASNYWLSLWTDPIVNGTQETHKVRSLV 1015  
 Qy 1019 YGALGILQGAIIFGYSMAVSIIGIFASRRHLHLLVNLRSFMSFFERTPSGNLVRFSK 1078  
 Db 1016 YGALGILQGAIIFGYSMAVSIIGIFASRRHLHLLVNLRSFMSFFERTPSGNLVRFSK 1075  
 Qy 1079 ELDTVDSMIPEVIMKMGSLFNVTGACIVILLATPTAAIIIPGLIYFFVQRFYVASSR 1138

[illegible]

```

RESULT 7
US-09-339-853A-83
; Sequence 83, Application US/09939853A
; Publication No. US20040039163A1
; GENERAL INFORMATION:
; APPLICANT: Burgess et al.
; TITLE OF INVENTION: No. US20040039163A1
; APPLICANT'S ATTORNEY: Same

```

```

; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939,853A
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/228,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 83
; LENGTH: 1527
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-939-853A-83

      45.4%; Score 4491.5; DB 11; Length 1527;
      Best Local Similarity 56.7%; Pred. No. 0;
      Matches 873; Conservative 260; Mismatches 371; Indels 35; Gaps 9

Qy      8  SAGSDPLMDNVTWNTSNPFTKCFQNTVLVWPCFYLMWACFPFYLYLSRHRDGYTQM 67
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      7  SGEIGSKFSDNLSVHTENPDLTFCQNSLLAWVPCYILWVALPCYLLYLRRHCRGYIL 66
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      68  TPLNKTKTALGCLFWLWVCWADLFYSFWSRSRGIFLAPVFLVSPDLLGTTLLATFLIQLE 127
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      67  SHLSKLRWLVGLVLCWCSWADLFVSFCHLVHGRAPAPVFVFTPLVGVGTMLLATLIQVE 126

Qy      128  RRKGVSQSGIMLTFWLVALVCALAILRSKIMTALKEDAQVDLFRDITFYVYFSLLIQVL 187
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      127  RLQGVQSSGVLLIIFWFLCVCAIIVPFRSKILLAKAEGEISDPFRFTFYIHFALVLSALI 186
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```







Db	1365	PODPILPSGTLRNMLDPFGRYSDIEDIWRTELSHL\$AFVSQPTGLDFQCSGEGDNL\$VG	1426
Qy	1434	QRQLVCLARALLRKTKILVLDEATAA\$VDLETDLLIQSTIR\$TQFEDCTVLTIAHRLNTIMD	1493
Db	1425	QRQLVCLARALLRKSRVLDEATAA\$DLETDLLIQSTIR\$TQFEDCTVLTIAHRLNTIMD	1484
Qy	1494	YTRVIVLDKGEIOEYGA\$PSDILLOQRGLFY\$MAKDAGL	1530
Db	1485	YNRVIVLDKGVAAEFD\$P\$VNLIAGGIFY\$GMAKDAGL	1521
RESULT 11			
US-10-667-891-1			
; Sequence 1, Application US/10667891			
; Publication No. US20040171024A1			
; GENERAL INFORMATION:			
; APPLICANT: ROTH, CHARLES W.			
; APPLICANT: BREY, PAUL T.			
; APPLICANT: HOLM, INGE			
; APPLICANT: GRAILLES, MARINE			
; APPLICANT: RZHETSKY, ANDREY			
; TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS IN DROSOPHILA AND			
; TITLE OF INVENTION: ANOPHELES			
; FILE REFERENCE: 03495.0294-00000			
; CURRENT APPLICATION NUMBER: US/10/667,891			
; CURRENT FILING DATE: 2003-09-23			
; PRIOR APPLICATION NUMBER: 60/413,469			
; PRIOR FILING DATE: 2002-09-26			
; NUMBER OF SEQ ID NOS: 76			
; SOFTWARE: PatentIn Ver. 3.2			
; SEQ ID NO 1			
; LENGTH: 1548			
; TYPE: PRT			
; ORGANISM: Drosophila melanogaster			
US-10-667-891-1			
Query Match 39.0%; Score 3860.5; DB 16; Length 1548;			
Best Local Similarity 50.0%; Pred. No. 2e-314;			
Matches 784; Conservative 274; Mismatches 438; Indels 71; Gaps 18;			
Qy	6	FCSADGSDPLWDN\$VNTNTSPD\$TKFQNTLV\$VVCFLY\$WA--CP\$FYFLYLSR\$H\$RG	63
Db	12	FCGS-----TF\$NATET\$YTD\$D\$FTPC\$EQAL\$VWTP\$CAF\$W\$AFV\$IDF\$Y\$VKAS\$L-DRN	66
Qy	64	YIQMTPLN\$KTKTAL\$G\$LLI\$W\$VC\$ADL\$F\$Y\$S\$F\$W\$S\$RG\$IF\$AP\$V\$F---LV\$S\$PTLL\$G\$IT\$T\$LLA	120
Db	67	-IP\$N\$KLN\$V\$K\$AL\$V\$N\$G\$LL\$V\$IT\$AL\$D\$L\$MAL\$V\$K\$G-GD\$E\$LP\$LY\$D\$L\$DV\$G\$P\$II\$K\$E\$AT\$F\$LL	124
Qy	121	TFLIQLERRKGV\$Q\$SG\$IM\$L\$FWL\$VAL\$V\$CAL\$AIL\$RSKI-----MT\$AL\$K\$E\$D\$A\$Q\$VD	168
Db	125	PIFIPLNRKY\$Q\$V\$QT\$G\$Q\$F\$IF\$W\$LL\$TVLS\$IPR\$CR\$TE\$V\$RL\$DA\$E\$R\$K\$IL\$N\$S\$Q\$P\$S\$EQ\$D\$F\$S\$W\$E	184
Qy	169	LFRDITFV\$Y\$F\$S\$LL\$II\$Q\$LV\$L\$S\$CF\$S\$DR\$S\$PL\$F\$E\$TH\$D\$N\$P\$C\$P\$E\$S\$AS\$F\$LS\$RT\$F\$W\$IT\$GLI	228
Db	185	EYQV\$S\$P\$F\$IF\$T\$F\$TS\$IM\$L\$N\$C\$F\$AD\$G\$P\$R\$Q\$TKY\$O\$R\$G\$E\$N\$E\$I\$P\$E\$L\$S\$AS\$F\$LS\$R\$IT\$Y\$Q\$W\$F\$D\$K\$M\$A	244
Qy	229	VRGYRQ\$P\$LG\$S\$DL\$W\$SL\$N\$K\$E\$D\$TS\$E\$Q\$V\$P\$V\$LV\$KN\$W\$K\$E\$C\$AK\$TK\$Q\$P\$K\$V\$V\$Y\$S\$K\$D\$P\$AQ\$P\$K\$E\$S	288
Db	245	LKGYN\$P\$LE\$E\$K\$DL\$W\$DL\$RPQ\$D\$S\$E\$W\$P\$IF\$A\$H\$W\$N\$Q\$V\$N\$K\$Y\$N\$K\$AR\$V-----E\$P\$K\$AQ\$F\$S\$NG	300
Qy	289	SKVDAN\$E\$VE\$A\$L\$IK\$S\$P\$Q\$E\$W\$N\$S\$P\$L\$K\$V\$LY\$K\$T\$F\$G\$PY\$F\$LM\$S\$F\$F\$F\$K\$AI\$H\$D\$LM\$M\$F\$S\$Q\$IL\$KL	348
Db	301	NVTE\$N\$PH\$G\$E-----K\$N\$G\$R\$K\$G\$M\$A\$S\$IM\$P\$PI\$Y\$K\$S\$F\$G\$V\$F\$G\$AL\$M\$K\$L\$T\$D\$T\$L\$T\$F\$AQ\$P\$Q\$V\$LSL	356
Qy	349	L\$IK\$F\$V\$N\$D\$TKA-PD\$N\$Q\$G\$Y\$F\$Y\$T\$V\$LL\$F\$V\$T\$AC\$L\$Q\$T\$LV\$L\$H\$O\$Y\$F\$H\$IC\$F\$V\$S\$G\$M\$R\$IK\$T\$AV\$IC\$N\$V\$R\$K\$A	407
Db	357	I\$IS\$F\$V\$E\$A\$Q\$AD\$E\$P\$E\$W\$K\$G\$IL\$Y\$V\$LL\$F\$V\$L\$A\$A\$Q\$T\$F\$IL\$G\$Y\$F\$H\$R\$M\$F\$IV\$G\$LR\$IR\$T\$AL\$IN\$AI\$Y\$R\$K\$A	416
Qy	408	LV\$IT\$N\$A\$R\$K\$S\$T\$V\$G\$E\$IV\$N\$M\$S\$V\$D\$A\$Q\$R\$E\$M\$D\$LAT\$Y\$IN\$M\$I\$W\$S\$AP\$L\$Q\$V\$L\$AL\$Y\$LL\$W\$L\$N\$G\$S\$V\$L	467
Db	417	LRI\$N\$S\$TK\$E\$S\$T\$V\$G\$E\$IV\$N\$M\$A\$V\$D\$A\$Q\$R\$E\$M\$LT\$Y\$IN\$M\$I\$W\$S\$AP\$L\$Q\$V\$L\$AL\$Y\$LL\$W\$L\$N\$G\$S\$V\$L	476
Qy	468	AG\$V\$A\$M\$V\$M\$V\$P\$V\$N\$A\$V\$M\$A\$K\$T\$K\$TYQ\$V\$A\$H\$M\$K\$S\$D\$N\$R\$IK\$M\$N\$E\$T\$ING\$IK\$V\$L\$K\$Y\$A\$W\$E\$L\$A\$P\$K\$D\$K	527

RESULT 13  
US-10-667-891-3  
; Sequence 3, Application US/10667891  
; Publication No. US20040171024A1  
; GENERAL INFORMATION:

APPLICANT: ROTH, CHARLES W.  
 APPLICANT: BREY, PAUL T.  
 APPLICANT: HOLM, INGE  
 APPLICANT: GRAILLES, MARINE  
 APPLICANT: EZHETSKY, ANDREY  
 TITLE OF INVENTION: MUTIDRUG RESISTANCE PROTEINS IN DROSOPHILA AND  
 TITLE OF INVENTION: ANOPHELES  
 FILE REFERENCE: 03495.0294-00000  
 CURRENT APPLICATION NUMBER: US/10/667,891  
 CURRENT FILING DATE: 2003-09-23  
 PRIOR APPLICATION NUMBER: 60/413,469  
 PRIOR FILING DATE: 2002-09-26  
 NUMBER OF SEQ ID NOS: 76  
 SOFTWARE: Patent in Ver. 3.2  
 SEQ ID NO 3  
 LENGTH: 1499  
 TYPE: PRT  
 ORGANISM: Anopheles gambiae  
 us-10-667-891-3

Query Match 36.4%; Score 3601.5; DB 16; Length 1499;  
 Best Local Similarity 47.9%; Pred. No. 1.2e-292;  
 Matches 744; Conservative 283; Mismatches 448; Indels 79; Gaps 20;

QY 1 MALRGCSADGSDPLMDMNVNTSNPDTKCFQNTVLVWPCFYLWACFPFFYLYLSRH 60  
 DB 1 MTFEEFCGG----PFWDDDLWREEDPDLTFCFQVILQWTCFPLFVFSMVEVLRIVTS 56  
 QY 61 DRGYIOMTLENKTKTALGELLWVWADLFYFWSERSGIFLAPVLPVSPITLIGIT---- 116  
 DB 57 RYRIDPWNFNITKMTFTFALVMWSVDL-----GVGLANDWIES--LLSLMPSLP 105  
 QY .117 -TLLATFLQLERKGVQSGQIMLTFWALVALCALAILRSKMT--ALKEDAQVLD--PRD 172  
 DB 106 FOIMAMALVFFRYKYGIRGTGTFIWFILKAFGIQMTKTEAMLDHVRSGTGDPAEPQF 165  
 QY 173 ITFYVYFSLLLIQLVLSCFSDRSPLESTHDPNCPSESSAFLSHITFWITGLIVRGY 232  
 DB 166 VSYTIQYVCCVLLLELFPDKREPRYSEWAKLNPNPDLRSFSSFLYLYPDSVAMRGF 225  
 QY 233 RQPLEGSDLSLNKETSTQVVPVLVWKKCEKATKRPVKVYSSKDPQPKSSKVD 292  
 DB 226 RKPLTDDMDYLDNPDTSALVPPDPKYYESVEKGRKQI-----AADKKAGKTN 276  
 QY 293 ANEEVEALVKSPOKEWNSLPKVLKTF--GPVFLMSFFFKATHDLMPFSGPOLKLLIK 351  
 DB 277 -----LVYKNAATNGSVLPAMVKA YGCFWFAQMLQFALSGL-QFASPYLMQEIWA 327  
 QY 352 FVNDTKAPDQGYFTVLLFTVLAQLTLVLHQYFHCYFSGMRKTAIVIGAVYRKALVIT 411  
 DB 328 VI-ALDGPFWKGMIIITGLFLSLIALFNGQYFHTFLVGFRIQTGLISAIYRKALRIS 386  
 QY 412 NSARKSTVGEIWNLSVDAQRPMDLATYINMWSAPLOVILAYLLWNLGSPVLAVGA 471  
 DB 387 SFAKDQTVGEIYNLMAVDAQRFELTSYLHWSAPLIATLCIYLLYELLGPAVPAIG 446  
 QY 472 VVVLAVPVNAVMAWTKTKTVQVAHMSKDNRIKLWNEILANGIKVLKYAWELAFKDKVLAI 531  
 DB 447 VVWIMIPITGFIATRDIDQVEQMKIKDERVKQKNEILGKIKLYAWEPFQDTIVTVV 506  
 QY 532 ROBELKVLKKSAYSALVAGTFTWYCTPFLVALCTFAVYVITDENNILDQAQTAFLVSLAFNI 591  
 DB 507 RNEELDVLSAAVYAGTFTVWMAFLVTLASFAVYVIMDEENVLDPQTAFLVALAFNI 566  
 QY 592 LRPLNIPMVISSIVQASVSLKRLRIFLSHELEPDSIERRPVKDGCGTNSITVFNATP 651  
 DB 567 LRPLAMFPMITFAMQAWYSIKRIDKFMNSELDENNVTNTHK-----SENALVEYKDTGF 621  
 QY 652 TWARSDDPTLNGITFSIPGALVAVGVQCGCKSSLLSALLAEMDKVEGHVAKGSAVY 711  
 DB 622 SWG-DDAPTLKLNILNRGKUSAVVGGVGTGKSSLLSALLGEMKMGKSVNTDGSIAV 680  
 QY 712 PQQAWIQNSLRNIFGCGLEBPYYSRVIQACALLPDLEILPSPGDRTEIGKGVNLSGG 771

DB 681 PQQAWIQNATLRDNLFRPFDQAKYDKVIECCALRPDLEMLPGDGTTEIGEKGINLSGG 740  
 QY 772 OKORVSLARAVYSNADIYLFDDPLSAVDAHVGKHFENVIKPKMLKKNKTRILVTHSMYSY 831  
 DB 741 QKQVALLARAVYADSEVYLFDDPLSAVDAHVGKHFENVIKPKMLVGRSLLVTHGTSF 800  
 QY 832 LPQDVIIIVMSGGKISEMGSYOELLARDGAFABFRTYASTEQEOQDAENGVTGSGPK 891  
 DB 801 LPFVEEIFVMKDGESGSGYOELLDDQKGAFAEFLTQHI---QEMDDEDE-----D 848  
 QY 892 EAKQWENGMLVYDSAGKQIQRLSSSSSYSGDI-----SRHNSHSTAELOKAE 938  
 DB 849 ELKIQEA-LKDGEAKTVQRAMSTRSQSSGNSGVRKRVSRAESRNSKPRAVEQTV 907  
 QY 939 AKKEETWKLMEADKAQTKQVLSVYDWMYKATGLFISFLSIFELFMCNHVSASNASYWSL 998  
 DB 908 AQSSA-TLIEKESATGAVGVVVIKFKGIGLWLGWSIFFSVINQOASIVANIWLTD 966  
 QY 999 WTDDPIVNGTQEHKTVRLSVYGAIGISQGIAGVFGSMASVIGIILASRCLHVDLHSLIR 1058  
 DB 967 WSEDPAAATDPSVRDWMYLVGGIGGAQSIALLIASVTLLALGCIKAAARELHNNLESMSR 1026  
 QY 1059 SPMSFFERTPSGNLVNRSKELDTVDSMIPEVVKPMGSLFNVIACIVILLATPAAII 1118  
 DB 1027 MPMSFPDTPPLGRIMNRFKSDVVDVNDILPQSIKAWLLMFFNVIGVVFVIGISTPFLAV 1086  
 QY 1119 IPLGLIYFFQVRFVYASSRQLKLESVRSRVPVYSHFNETLLGVSVIRAFESQERIHOS 1178  
 DB 1087 VPAFVLIYLIQKFIATSRQKRLSVTRSPFIYSHFGESITGQSTIRAYGQDRPMNES 1146  
 QY 1179 DLKVDENQKAYPSYIVANRWLAVRLECVGNCIVLFAALFAVISHRSLSAGLVGSYSYL 1238  
 DB 1147 EQRDVYNQLTSPYSIIANRWLAVRLELVCALVFAALFAVARDTIGQATVGLSISYAL 1206  
 QY 1239 QUTTVLWMLVRMSSEMTNIVAVERLKEYSETEKAPWQIQETAPPSPQVGRVFRNY 1298  
 DB 1207 QISATSLFLVRMTAEVETNVAIERLEEYTVLPREAWQLGHV--DKAMPVEKGVFKDY 1264  
 QY 1299 CLURYREDLDFVLRLHNVITNGEKGIVGRTGAGKSSLTGLGLFRINESAGEIIIDGINI 1358  
 DB 1265 QIRYREGDLVIRGSLNVRGGEKIGVGTGAGKSSLTGLGLFRIVEAAGQIIDDGLDI 1324  
 QY 1359 AKIGHDLRFKTIITIPQDPVLFSGSLRWMLDPFQSYDSEVWTSLELAHLKDFVSALPK 1418  
 DB 1325 SKWGLHQLRGLRTIIPQDPVLFSGTLRANVDPFKSYDSDLVWKALELSHLKTFVKGLAAG 1384  
 QY 1419 LDHECAEGENLSVQORQLVCLARALLKTKILVLDATAAVDLETDLLIQSTIRTFQED 1478  
 DB 1385 LDHEIAENGENLSVQORQLICLARAVLRKTKVILDEATAAVDLETDLLIQSTIRTFAD 1444  
 QY 1479 CTVLTAHRLNTIMDYTRVIVLDKGEIOEYGAPSDLOOR-GLFYSMADKAGLV 1531  
 DB 1445 CTVLTAHRLNTILSDRVLDKGLVAECDSFQNLNANRESIFFGMKNAGIV 1498

RESULT 14  
 US-10-363-112-4  
 ; Sequence 4, Application US/10363112  
 ; Publication No. US20040091964A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: THE AUSTRALIAN NATIONAL UNIVERSITY  
 ; APPLICANT: BOARD, PHILLIP  
 ; APPLICANT: HARRIS, MATTHEW  
 ; TITLE OF INVENTION: MODIFIED PROTEINS, ISOLATED NOVEL PEPTIDES, AND USES THEREOF  
 ; FILE REFERENCE: 007643-0302189  
 ; CURRENT APPLICATION NUMBER: US/10/363,112  
 ; CURRENT FILING DATE: 2003-11-03  
 ; PRIOR APPLICATION NUMBER: PCT/AU01/01093  
 ; PRIOR FILING DATE: 2001-08-31  
 ; PRIOR APPLICATION NUMBER: 60/229,663  
 ; PRIOR FILING DATE: 2000-08-31  
 ; NUMBER OF SEQ ID NOS: 62





THIS PAGE BLANK (USPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2005, 10:57:19 ; Search time 138.282 Seconds  
(without alignments)

5389.603 Million cell updates/sec

Title: US-10-665-283-1

Perfect score: 9903

Sequence: 1 MALRGFCADGSDPLMDNV.....RSVAVAKPKFSPDLS 1927

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

- 1: Geneseq1980s:\*
- 2: Geneseq1990s:\*
- 3: Geneseq2000s:\*
- 4: Geneseq2001s:\*
- 5: Geneseq2002s:\*
- 6: Geneseq2003as:\*
- 7: Geneseq2003bs:\*
- 8: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7860	79.4	1531	2 AAW57486	AAW57486 Human MRP
2	7860	79.4	1531	2 AAW74471	Aaw74471 Human mul
3	7860	79.4	1531	2 AAW99894	Aaw99894 Human mul
4	7860	79.4	1531	3 AAY55799	Aay55799 Human mul
5	7860	79.4	1531	3 AAY78873	Aay78873 Multidrug
6	7860	79.4	1531	5 ABG61810	ABg61810 Prostate
7	7860	79.4	1531	6 ABM35012	ABm35012 Cancer ba
8	7860	79.4	1531	6 ADB20865	ADB20865 MRP1 base
9	7860	79.4	1531	7 ADB87954	ADB87954 Human UGT
10	7860	79.4	1531	7 ADB96937	ADB96937 Human MDR
11	7860	79.4	1531	7 ADB92128	ADB92128 Human MDR
12	7860	79.4	1531	7 ADD44764	Add44764 Human Pro
13	7860	79.4	1531	8 ADN97111	Adn97111 MRP1 prot
14	7856	79.3	1530	7 ADN95929	Adn95929 Human BEC
15	7849	79.3	1531	2 AAR54928	Aar54928 Human mul
16	7849	79.3	1531	2 AAR93153	Aar93153 Multidru
17	7849	79.3	1531	2 AAW57485	AAW57485 Human mul
18	7849	79.3	1531	2 AAW74470	Aaw74470 Human mul
19	7849	79.3	1531	2 AAW99893	Aaw99893 Human mul
20	7849	79.3	1531	3 AAY55798	Aay55798 Human mul
21	7849	79.3	1531	3 AAY78872	Aay78872 Human mul
22	7849	79.3	1531	3 AAB03582	Aab03582 Multidrug
23	7810	76.8	1489	2 AAR96952	Aar96952 Multidru
24	7587.5	76.6	1482	8 ADN03902	Adn03902 Antipsori
25	7587.5	76.6	1482	8 ADP23422	Adp23422 PRO polyp

26	7572	76.5	1481	5 ABP52108	Abp52108 Homo sapi
27	7560.5	76.3	1482	2 AAR96953	Aar96953 Multi-dru
28	7500.5	75.7	1472	2 AAR96954	Aar96954 Multi-dru
29	7230.5	73.0	1530	7 ADF56611	Adf56611 Bovine MR
30	7194	72.6	1417	2 AAR96955	Aar96955 Multi-dru
31	7037	71.1	1388	7 ADJ69912	Adj69912 Human hea
32	7002.5	70.7	1528	2 AAW57487	Aaw57487 Murine mu
33	7002.5	70.7	1528	2 AAW74472	Aaw74472 Mouse mul
34	7002.5	70.7	1528	2 AAW99895	Aaw99895 Mouse mul
35	7002.5	70.7	1528	3 AAY55800	Aay55800 Murine mu
36	7002.5	70.7	1528	3 AAY78874	Aay78874 Murine mu
37	6916	69.8	1372	2 AAR98211	Aar98211 Multi-dru
38	6911	69.8	1345	2 AAR96956	Aar96956 Multi-dru
39	4491.5	45.4	1527	5 AAU91309	Aau91309 Human pro
40	4491.5	45.4	1527	5 ABP52110	Abp52110 Homo sapi
41	4491.5	45.4	1527	6 ABR58639	Abt58639 Human can
42	4491.5	45.4	1527	7 ADE54387	Adt54387 Human pro
43	4491.5	45.4	1527	7 ADN40024	Adn40024 Cancer/an
44	4491.5	45.4	1527	8 ADP24088	Adp24088 PRO polyp
45	4487.5	45.3	1527	2 AAY43543	Aay43543 A human M

#### ALIGNMENTS

##### RESULT 1

AAW57486  
ID AAW57486 standard; protein; 1531 AA.

XX AAW57486;

DT 14-AUG-1998 (first entry)

DE Human MRP variant ltPgpa (Lei/Pgpa).

XX Multidrug resistance-associated protein; MRP; tumour; human; variant;  
KW multidrug resistance; MDR; leishmania P-glycoprotein; ltPgpa; Lei/Pgpa.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 685 /label= L685S

FT /note= "wild-type Leu is replaced by Ser"

FT Misc-difference 1282

FT /label= R1282A

FT /note= "wild-type Arg is replaced by Ala"

XX US5766880-A.

XX 16-JUN-1998.

XX 05-JUN-1995; 95US-00463092.

XX 27-OCT-1992; 92US-00966923.

XX 08-MAR-1993; 93US-00029340.

XX 26-OCT-1993; 93US-00141893.

XX 20-MAR-1995; 95US-00407207.

XX (TOOH ) UNIV QUEENS KINGSTON.

XX Cole SP, Deeley RG;

XX WPI; 1998-361687/31.

XX N-FSDB; AAV31498.

XX DNA encoding protein associated with multi-drug resistance - useful for  
PT as probe for identifying multi-drug resistant tumour cells.

XX Claim 1; Col 67-78; 82pp; English.

XX This represents a variant of the human multidrug resistance-associated  
CC protein (MRP). This natural variant is a leishmania P-glycoprotein related

molecule ltpgpa (Lei/PgpA). The human and murine MRP nucleic acid molecules can be used as probes for identifying multidrug resistant tumour cells by hybridisation to mRNA from tumour cells. The antisense nucleic acid can be used to reverse multidrug resistance (MDR). A recombinant expression vector containing the MRP nucleic acid molecules operatively linked to at least one regulatory sequence can be used to transform a host cell to produce a recombinant MDR-associated protein

Sequence 1531 AA;

Query Match 79.4%; Score 7860; DB 2; Length 1531;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CC 1 MAIRGFCSDGSDPLMDNNTWNTSNPDTKCFQNTVLVWPCFYLWACFPFYFLYLSRH 60  
 CC |  
 CC 1 MAIRGFCSDGSDPLMDNNTWNTSNPDTKCFQNTVLVWPCFYLWACFPFYFLYLSRH 60  
 CC |  
 CC 61 DRGYIQMTPLNKTALGFLWVWADLFYFWSRSGIFLAPVPLVSPPTLLGTTLLA 120  
 CC |  
 CC 61 DRGYIQMTPLNKTALGFLWVWADLFYFWSRSGIFLAPVPLVSPPTLLGTTLLA 120  
 CC |  
 CC 121 TFLIOLERRKGVQSSGIMLTFWLVAVCALAIRSKIMTALKEDAQVDLFRDITFYVYFS 180  
 CC |  
 CC 121 TFLIOLERRKGVQSSGIMLTFWLVAVCALAIRSKIMTALKEDAQVDLFRDITFYVYFS 180  
 CC |  
 CC 181 LLLIOLVLSFSDRPLSETIHDNPNCPSSASFLSRITFWITGLIVRGYRQPLEGSD 240  
 CC |  
 CC 181 LLLIOLVLSFSDRPLSETIHDNPNCPSSASFLSRITFWITGLIVRGYRQPLEGSD 240  
 CC |  
 CC 241 LWSLNKEDTSEQVPLVKNWKECAKTRQKPVVYSSKDPAPKESKVDANEVEAL 300  
 CC |  
 CC 241 LWSLNKEDTSEQVPLVKNWKECAKTRQKPVVYSSKDPAPKESKVDANEVEAL 300  
 CC |  
 CC 301 IVKSPQKWNPSLFVLYNTGPGYFLMGFFFKAIHDLMMFSGPQILKLLIKFVNDTKAPD 360  
 CC |  
 CC 301 IVKSPQKWNPSLFVLYNTGPGYFLMGFFFKAIHDLMMFSGPQILKLLIKFVNDTKAPD 360  
 CC |  
 CC 361 WQGYEYTVLLFVTAQLTILVHOYPHICFVSGMRIKTAIVGAVYRKALVITNSARKSTV 420  
 CC |  
 CC 361 WQGYEYTVLLFVTAQLTILVHOYPHICFVSGMRIKTAIVGAVYRKALVITNSARKSTV 420  
 CC |  
 CC 421 GEIVNLSVDAQRFMDLTYINWISAPLOVTLALYLLWNLGSPVLGAVVWMLVPVN 480  
 CC |  
 CC 421 GEIVNLSVDAQRFMDLTYINWISAPLOVTLALYLLWNLGSPVLGAVVWMLVPVN 480  
 CC |  
 CC 481 AVAMKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAWELAFKDKVLAIROBELKVLK 540  
 CC |  
 CC 481 AVAMKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAWELAFKDKVLAIROBELKVLK 540  
 CC |  
 CC 541 KSAYLSAVGTFTWCTPFLVALCTFAVVTIDENNILDAQTAFAVSLALFNILRPPLNLP 600  
 CC |  
 CC 541 KSAYLSAVGTFTWCTPFLVALCTFAVVTIDENNILDAQTAFAVSLALFNILRPPLNLP 600  
 CC |  
 CC 601 MWISSIVQASVSLKELRIFLSHEELEPDSIERPVPKGGGINSITVRNATFTWARSDDPT 660  
 CC |  
 CC 601 MWISSIVQASVSLKELRIFLSHEELEPDSIERPVPKGGGINSITVRNATFTWARSDDPT 660  
 CC |  
 CC 661 LINGITFSIPEGALVAVVQGVCGKSSLLSALLAEMDKVEGHVAIKGSVAVYVPPQAWIQND 720  
 CC |  
 CC 661 LINGITFSIPEGALVAVVQGVCGKSSLLSALLAEMDKVEGHVAIKGSVAVYVPPQAWIQND 720  
 CC |  
 CC 721 SIRENILFGCOLLEPYYSVTOACALLPDLELTPSGDRTEIGKGVNLSSGQKQVSLAR 780  
 CC |  
 CC 721 SIRENILFGCOLLEPYYSVTOACALLPDLELTPSGDRTEIGKGVNLSSGQKQVSLAR 780  
 CC |  
 CC 781 AVYSNADILYLFDDPLSADVAHVGHIFENVIGPKGMLKNKTRILVTHSMVSLPOVDVIV 840  
 CC |  
 CC 781 AVYSNADILYLFDDPLSADVAHVGHIFENVIGPKGMLKNKTRILVTHSMVSLPOVDVIV 840  
 CC |  
 CC 841 MSGGKISEMGVQELLARDGAFAPFLRYASTEQSDAEENGVTGVSFGPKGKQKMGNGM 900  
 CC |  
 CC 841 MSGGKISEMGVQELLARDGAFAPFLRYASTEQSDAEENGVTGVSFGPKGKQKMGNGM 900  
 CC |

QY 901 LVTDAGKQLOLQSSSSSYSGDISRHNHSTAELOKABAKKEETWKLMEADKAQGTQVKL 960  
 Db |  
 QY 901 LVTDAGKQLOLQSSSSSYSGDISRHNHSTAELOKABAKKEETWKLMEADKAQGTQVKL 960  
 Db |  
 QY 961 SVYDYMKAIGLFTSIFLFCMCHVGSALASNYWLSLWTDPIVNGTQHTKVRLSVYG 1020  
 Db |  
 QY 961 SVYDYMKAIGLFTSIFLFCMCHVGSALASNYWLSLWTDPIVNGTQHTKVRLSVYG 1020  
 Db |  
 QY 1021 ALGISQGIAGFYGSMNAVSIIGGILASRCLHVDLLHSILSPMSFFERTSGNLVNRFSKEL 1080  
 Db |  
 QY 1021 ALGISQGIAGFYGSMNAVSIIGGILASRCLHVDLLHSILSPMSFFERTSGNLVNRFSKEL 1080  
 Db |  
 QY 1081 DTVDMSIPEVIMFPGSLFNIVGACIVILLATPAAIIIPPLGLIYFPVQPFYVASSRQL 1140  
 Db |  
 QY 1081 DTVDMSIPEVIMFPGSLFNIVGACIVILLATPAAIIIPPLGLIYFPVQPFYVASSRQL 1140  
 Db |  
 QY 1141 KRLESVSRSPVYSHFNETLLGVSVIRAFEEQERFIHQSDLVKVDENOKAYPSIVANRWLA 1200  
 Db |  
 QY 1141 KRLESVSRSPVYSHFNETLLGVSVIRAFEEQERFIHQSDLVKVDENOKAYPSIVANRWLA 1200  
 Db |  
 QY 1201 VRLECVCNIVLFAALFAVISHRSLSAGLVGSYSLSQVTTYLNWLVRMSSEMETNIVA 1260  
 Db |  
 QY 1201 VRLECVCNIVLFAALFAVISHRSLSAGLVGSYSLSQVTTYLNWLVRMSSEMETNIVA 1260  
 Db |  
 QY 1261 VERLKEYSETEKEAPWQIQTAPPSSWPQVGRVFRNYCLRYREDLDFVLRIHNVITNGG 1320  
 Db |  
 QY 1261 VERLKEYSETEKEAPWQIQTAPPSSWPQVGRVFRNYCLRYREDLDFVLRIHNVITNGG 1320  
 Db |  
 QY 1321 EKVIGVGTGAGKSLTLGLFRINESAGEIIDIIGINIAKIGLHDLRPFITIIPODPVLF 1380  
 Db |  
 QY 1321 EKVIGVGTGAGKSLTLGLFRINESAGEIIDIIGINIAKIGLHDLRPFITIIPODPVLF 1380  
 Db |  
 QY 1381 SGSLRMNLDPPSQYSDEEVWTSLELAHLKDFVSALPKDLHECAEGGNSLVGQRLVCL 1440  
 Db |  
 QY 1381 SGSLRMNLDPPSQYSDEEVWTSLELAHLKDFVSALPKDLHECAEGGNSLVGQRLVCL 1440  
 Db |  
 QY 1441 ABALLRKTILVDEATAAVALDLEDDLIQSTIRPTQFEDCTVLTIAHRLNTIMDYTRVIV 1500  
 Db |  
 QY 1441 ABALLRKTILVDEATAAVALDLEDDLIQSTIRPTQFEDCTVLTIAHRLNTIMDYTRVIV 1500  
 Db |  
 QY 1501 DKGIQIEYGAPSDLLQQRGLFYSMKADAGLV 1531  
 Db |  
 QY 1501 DKGIQIEYGAPSDLLQQRGLFYSMKADAGLV 1531  
 Db |

RESULT 2  
 AAW74471  
 ID AAW74471 standard; protein, 1531 AA.  
 XX AAW74471;  
 AC AAW74471;  
 XX 18-MAY-1999 (first entry)  
 DT Human multidrug resistance-associated protein variant.  
 XX Multidrug resistance-associated protein; MDR; human; diagnosis;  
 KW MDR tumour cell identification; cancer therapy.  
 XX Homo sapiens.  
 OS US5882875-A.  
 PN 16-MAR-1999.  
 PD 05-JUN-1995; 95US-00462109.  
 PF 27-OCT-1992; 92US-00966923.  
 PR 08-MAR-1993; 93US-00029340.  
 PR 26-OCT-1993; 93US-00141893.  
 PR 20-MAR-1995; 95US-00407207.  
 XX (TOOH ) UNIV QUEENS KINGSTON.  
 PA

XX	Cole SPC, Dealey RG;	Db		481	AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKLYAWELAFKDKVLAIROBELKVLK	540
PI						
DR	WPI: 1999-214061/18.	Qy		541	KSAYLSAVGTFTWCTPFLVALCTFAVVVTIDENNILDAQTAFAVSLALFNILRPPNLILP	600
DR	N-PSDB; AAX21977.					
XX		Db		541	KSAYLSAVGTFTWCTPFLVALCTFAVVVTIDENNILDAQTAFAVSLALFNILRPPNLILP	600
PT	Identifying a multidrug resistant tumor cell by contacting the cell with	Qy		601	MVSIISVQASVSLKRLRIFLSHEELEPDSIERRPVKOGGTSITVRNATFTWASDDPPT	660
PT	an antibody/antigen-binding fragment - which binds to an expressed					
PT	protein encoded by multidrug resistance-associated protein (MRP) nucleic	Db		601	MVSIISVQASVSLKRLRIFLSHEELEPDSIERRPVKOGGTSITVRNATFTWASDDPPT	660
XX						
PS	Claim 3; Col 69-80; 80pp; English.	Qy		661	LNGITFSIPEGALVAVVQVCGGKSSLLSALLAEMDKVEGHVAKGSVAVYVPOQAWIQND	720
XX						
CC	This sequence is the human multidrug resistance-associated (MDR) protein.	Db		661	LNGITFSIPEGALVAVVQVCGGKSSLLSALLAEMDKVEGHVAKGSVAVYVPOQAWIQND	720
CC	The invention relates to a method for identifying a multidrug resistant	Qy		721	SIRENIFGCOLPEPYRVSVOACALLPDLBILPSGDRTEIGEGKVNLSGGOKQVSLAR	780
CC	(MDR) tumour cell. Compositions and methods utilising the MDR proteins					
CC	can be used to treat patients with tumours displaying multidrug	Db		721	SIRENIFGCOLPEPYRVSVOACALLPDLBILPSGDRTEIGEGKVNLSGGOKQVSLAR	780
CC	resistance, particularly those displaying resistance to anthracyclines,	Qy		781	AVYGNADIYLPDDPLSAVDAAHVGHKHPENVIGPKGMLKNKTRILVTHSMSYLPQVDVIV	840
CC	epidophyllotoxins, vinca alkaloids, and hydrophobic drugs. The methods					
CC	for inhibiting/killing a MDR tumour cell can be useful for treating	Db		781	AVYGNADIYLPDDPLSAVDAAHVGHKHPENVIGPKGMLKNKTRILVTHSMSYLPQVDVIV	840
CC	breast cancer, leukaemias, fibrosarcomas, cervical cancer, gliomas,	Qy		841	MSGKISEMSGYQELLARDGAFAPFLRTYASTEQEODAEENGVTGVSQPGKEAKQMGNG	900
CC	thyomas, neuroblastomas and lung cancer. The MDR DNA sequences when					
CC	labeled are useful as molecular probes for diagnosing multidrug	Db		841	MSGKISEMSGYQELLARDGAFAPFLRTYASTEQEODAEENGVTGVSQPGKEAKQMGNG	900
CC	resistance of a tumour (using cells from a tumour biopsy) and for	Qy		901	LVTDSAGKOLQORQLSSSSSYSGDISRRHNSFAELQKAEKKEETWKLMEADKAQTQVKL	960
CC	designing ribozymes which are capable of cleaving a single-stranded					
CC	nucleic acid encoding a protein having MRP activity. Recombinant	Db		901	LVTDSAGKOLQORQLSSSSSYSGDISRRHNSFAELQKAEKKEETWKLMEADKAQTQVKL	960
CC	expression vectors containing human MDR coding sequences can be	Qy		961	SVYNDYKKAIGLFIISFLSIFLMCNHVSALASNYWLSLWTDPIVNGTQEHTKVRLSVYG	1020
CC	transfected into a drug sensitive cell line to produce a protein in the					
CC	cell which confers MDR, protecting non-resistant non-tumour cells from	Db		961	SVYNDYKKAIGLFIISFLSIFLMCNHVSALASNYWLSLWTDPIVNGTQEHTKVRLSVYG	1020
CC	the effects of chemotherapeutics has major clinical importance. Cells	Qy		1021	ALGISQGIATVFGYMAVSGIGLILASRCLHVDLLHSILRSPSPFBERTSGNLRVPSKEL	1080
CC	transformed with the MDR coding sequences are useful for testing					
CC	potential therapeutic agents for their effectiveness against MDR cells	Db		1021	ALGISQGIATVFGYMAVSGIGLILASRCLHVDLLHSILRSPSPFBERTSGNLRVPSKEL	1080
CC	and for identifying chemosensitisers of a therapeutic agent	Qy		1081	DTVDSMIPEVTKPMFMSLFNIVGACIVILLATPAAIIIPPLGLIYFFVQRPVYASSRQL	1140
XX						
SQ	Sequence 1531 AA;	Db		1081	DTVDSMIPEVTKPMFMSLFNIVGACIVILLATPAAIIIPPLGLIYFFVQRPVYASSRQL	1140
		Qy		1141	KRLSVSRSPVYSHFNETLLGVSVIRAFEEQERFIHQSDLVKVDENQKAYPSIVANRWLA	1200
		Db		1141	KRLSVSRSPVYSHFNETLLGVSVIRAFEEQERFIHQSDLVKVDENQKAYPSIVANRWLA	1200
		Qy		1201	VRLECVGNCIVLFAALFAVISRHSLSAGLVGSYSLOVTTYLNLVLRMSSEMETNIVA	1260
		Db		1201	VRLECVGNCIVLFAALFAVISRHSLSAGLVGSYSLOVTTYLNLVLRMSSEMETNIVA	1260
		Qy		1261	VERLKEYSETEKAPWQIOETAPPSSWPQVGRVFRNYCLARYREDLDFVLRHINVTINGG	1320
		Db		1261	VERLKEYSETEKAPWQIOETAPPSSWPQVGRVFRNYCLARYREDLDFVLRHINVTINGG	1320
		Qy		1321	EKVGIWGTGAGKSSLTGLFRINESABEIIIDGINIAKIGLHDLRKFITIIPODPVLF	1380
		Db		1321	EKVGIWGTGAGKSSLTGLFRINESABEIIIDGINIAKIGLHDLRKFITIIPODPVLF	1380
		Qy		1381	SGSLRMLNLDPPSQVSDDEEVTSLAHLKDFVSALPDKLDHECAGGENLSVGQRLVCL	1440
		Db		1381	SGSLRMLNLDPPSQVSDDEEVTSLAHLKDFVSALPDKLDHECAGGENLSVGQRLVCL	1440
		Qy		1441	ARALLRKTIIIVLEATAVDETDLLIQTIRTFQEDCTVLTIAHRLNTIMDYTRVIVL	1500
		Db		1441	ARALLRKTIIIVLEATAVDETDLLIQTIRTFQEDCTVLTIAHRLNTIMDYTRVIVL	1500
		Qy		1501	DKGETQETGAPSDLLIQORGLFYSMADAGLV	1531
		Db		1501	DKGETQETGAPSDLLIQORGLFYSMADAGLV	1531

AAW99894	XX	AAW99894 standard; protein; 1531 AA.	Db	241	LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPKVYSSKDPAPQKESKVDANEVEAL	300
AC	XX	AAW99894;	Qy	301	IVKSPQKEWNPSPFKVLYKTFPGPYPLMSFFPKAIHDLMMFSGPQLKLLIKFVNDTKAPD	360
XX	DT	10-JUN-1999 (first entry)	Db	301	IVKSPQKEWNPSPFKVLYKTFPGPYPLMSFFPKAIHDLMMFSGPQLKLLIKFVNDTKAPD	360
DE	XX	Human multidrug resistance-associated protein natural variant.	Qy	361	WQGFYTVLLFVTAQLOFLVHOFPHICFVSGMRIKTAVIGAVYRKALVITNSARKSTV	420
XX	KW	Human; multidrug resistance-associated protein; MRP; cytotoxic drug;	Db	361	WQGFYTVLLFVTAQLOFLVHOFPHICFVSGMRIKTAVIGAVYRKALVITNSARKSTV	420
KW	XX	Cancer; chemotherapy.	Qy	421	GEIVNLSVDAQRFRMDLATYINMIWSAPLQVILALYLNLGLSPVLAGVAVMLVMPVN	480
OS	XX	Homo sapiens.	Db	421	GEIVNLSVDAQRFRMDLATYINMIWSAPLQVILALYLNLGLSPVLAGVAVMLVMPVN	480
XX	PN	US5891724-A.	Qy	481	AVMAMKTTYQVAHMKSKDNRIKLMNEILNGIKVLKYAWELAFKDKVLAIROBELKVLK	540
XX	PD	06-APR-1999.	Db	481	AVMAMKTTYQVAHMKSKDNRIKLMNEILNGIKVLKYAWELAFKDKVLAIROBELKVLK	540
XX	XX	05-JUN-1995; 95US-00460907.	Qy	541	KSAYLSAVGTFTWCTPFLVALCTFAVYVTTDENNILDQAQTAFAVSLALFNILRPLNLP	600
PR	XX	27-OCT-1992; 92US-00966923.	Db	541	KSAYLSAVGTFTWCTPFLVALCTFAVYVTTDENNILDQAQTAFAVSLALFNILRPLNLP	600
PR	XX	08-MAR-1993; 93US-00029340.	Qy	601	MVSISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGSGTNSITVRNATFTWARSDEPT	660
PR	XX	26-OCT-1993; 93US-00141893.	Db	601	MVSISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGSGTNSITVRNATFTWARSDEPT	660
PR	XX	20-MAR-1995; 95US-00407207.	Qy	661	LANGITFSIPEGALVAVVGVQCGKSSLLSALLAEMDKVEGHVAIKGVYVPOQAWIQND	720
XX	PA	(TOOH ) UNIV QUEENS KINGSTON.	Db	661	LANGITFSIPEGALVAVVGVQCGKSSLLSALLAEMDKVEGHVAIKGVYVPOQAWIQND	720
XX	PI	Cole SPC, Deeley RG;	Qy	721	SLRENILFGCOLEBPYRVSIVQACALLPDLEILPSGDRTEIGEGVNLSSGQKORVSLAR	780
XX	XX	WPI; 1999-253868/21.	Db	721	SLRENILFGCOLEBPYRVSIVQACALLPDLEILPSGDRTEIGEGVNLSSGQKORVSLAR	780
DR	XX	N-PSDB; AAX19818.	Qy	781	AVYGNADIYLPDDPLSAVDAHVGHKIFENVIGPKGMLKNKTRILVTHSMSYLPQVDVIV	840
XX	PT	Protecting mammalian cells against cytotoxic drugs.	Db	781	AVYGNADIYLPDDPLSAVDAHVGHKIFENVIGPKGMLKNKTRILVTHSMSYLPQVDVIV	840
XX	BS	Claim 2; Col 79-86; 82pp; English.	Qy	841	MSGKISSEMGSYQELLARDGAFELRTYASTEQDAEENGVTGVSFGPGKEAKOMENG	900
CC	XX	The present sequence represents a human multidrug resistance-associated	Db	841	MSGKISSEMGSYQELLARDGAFELRTYASTEQDAEENGVTGVSFGPGKEAKOMENG	900
CC	CC	protein (MRP). The present invention also describes a method for	Qy	901	LVTDSAGKQLOQRLSSSSSYSGDLSRHNSHNSAELOKAEKBEETWKLMEADKAQTGVKL	960
CC	CC	protecting a mammalian cell against the cytotoxicity of anthracyclines,	Db	901	LVTDSAGKQLOQRLSSSSSYSGDLSRHNSHNSAELOKAEKBEETWKLMEADKAQTGVKL	960
CC	CC	epidophyllotoxins and Vinca alkaloids (A) by introducing into it a	Qy	961	SVYWDYKKAIGLFTSFLSIFLFCMCHVSALASNYWLSLWTDPIVNGTQHTKVRLSVYG	1020
CC	CC	nucleic acid (I) that hybridizes under stringent conditions to a nucleic	Db	961	SVYWDYKKAIGLFTSFLSIFLFCMCHVSALASNYWLSLWTDPIVNGTQHTKVRLSVYG	1020
CC	CC	acid (II) that encodes an MRP protein (III). Introduction of (I) protects	Qy	1021	ALGISQGIAGVFGYSMAVSIIGILASRCHVDLHLSILSRSPMSFFERTPSGNLVNRFSEL	1080
CC	CC	cells against cytotoxic effects of (A), particularly to protect normal	Db	1021	ALGISQGIAGVFGYSMAVSIIGILASRCHVDLHLSILSRSPMSFFERTPSGNLVNRFSEL	1080
CC	CC	cells against (A) being used for treatment of cancers. Cells transformed	Qy	1081	DTVDSMPIPEVTKMPGMSLFNVIGACIVILLATPIAAIIIPPLGLIYFFVQFYVASSRQL	1140
CC	CC	with (I) can be used to screen for agents that affect multidrug	Db	1081	DTVDSMPIPEVTKMPGMSLFNVIGACIVILLATPIAAIIIPPLGLIYFFVQFYVASSRQL	1140
CC	CC	resistance or are directly toxic to multidrug resistant cells, i.e.	Qy	1141	KRLSVSRSPVYSHFNETHLLGVSVIRAFEPQERFIHOSDLKVDENQKAYTPSVIVANRWLA	1200
CC	CC	potential therapeutics for multidrug-resistant cancers. Conferring	Db	1141	KRLSVSRSPVYSHFNETHLLGVSVIRAFEPQERFIHOSDLKVDENQKAYTPSVIVANRWLA	1200
CC	CC	resistance to normal cells should allow an increase in the dose of (A)	Qy	1201	VRLEVCNVCIVLFAALFAVLSRHSLSAGLVLSYSLSQVTTYLNWLVRSSSEMETNIVA	1260
CC	CC	that can be administered safely	Db	1201	VRLEVCNVCIVLFAALFAVLSRHSLSAGLVLSYSLSQVTTYLNWLVRSSSEMETNIVA	1260
XX	XX	Sequence 1531 AA;	Qy	1261	VERLKEYSETEKEAPWQIQETAPSSWPQVGRVFRNCLRYREDLDFVLRIHNTINGG	1320
Query Match		79.4%; Score 7860; DB 2; Length 1531;	Db	1321	EKVGIVCRTGAGKSSLTGLFRINESAGEIIIDGINIAKGLHDLRFKTIIPQDPVILF	1380
Best Local Similarity		100.0%; Pred. No. 0;	Qy			
Matches 1531; Conservative		0; Mismatches 0; Indels 0; Gaps 0;	Db			
Qy	1	MALRGFCSDGSDPLMDNNTWNTSNPDTKCFQNTLVVWPCFYLWACFPFFFLYLSRH	60			
Db	1	MALRGFCSDGSDPLMDNNTWNTSNPDTKCFQNTLVVWPCFYLWACFPFFFLYLSRH	60			
Qy	61	DRGYIQMTPLNKTALGFLWVWADLFYSFWERSRGIFLAPVFLVSPILLGITLLA	120			
Db	61	DRGYIQMTPLNKTALGFLWVWADLFYSFWERSRGIFLAPVFLVSPILLGITLLA	120			
Qy	121	TFLIQLERRKGVOSSGIMLTFWLVALCALAIRSKIMTALKEDAQVDLFRDITFYVFS	180			
Db	121	TFLIQLERRKGVOSSGIMLTFWLVALCALAIRSKIMTALKEDAQVDLFRDITFYVFS	180			
Qy	181	LLLIQLVLSFSDRSPFLSETIHDNPNCPSSASFLSRITFWITGLIVRGYRQPLEGSD	240			
Db	181	LLLIQLVLSFSDRSPFLSETIHDNPNCPSSASFLSRITFWITGLIVRGYRQPLEGSD	240			
Qy	241	LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPKVYSSKDPAPQKESKVDANEVEAL	300			

Db 1321 EVGIVGRTGAGKSLTLGLFRINSAEGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
Qy 1381 SGLSLRNLDPPFSQYSDDEEVTTSLELAHLKDFVSALPDKLDHECARGGENLSVGQRLVCL 1440  
Db 1381 SGLSLRNLDPPFSQYSDDEEVTTSLELAHLKDFVSALPDKLDHECARGGENLSVGQRLVCL 1440  
Qy 1441 APALLRKTKILVLDATAVADLETDLIQSTIRTQFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
Db 1441 APALLRKTKILVLDATAVADLETDLIQSTIRTQFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
Qy 1501 DKGEIQEYQVAGPSDLQORGLFYSMADAGLV 1531  
Db 1501 DKGEIQEYQVAGPSDLQORGLFYSMADAGLV 1531

RESULT 4

AA55799  
ID AA55799 standard; protein; 1531 AA.  
XX AC AA55799;  
XX DT 28-FEB-2000 (first entry)  
XX DE Human multidrug resistance-associated protein (MRP) variant.  
XX KW Chemosensitizer; multidrug resistance-associated protein; MRP; human;  
XX KW therapeutic agent; P-glycoprotein-mediated multidrug resistance; lung;  
XX KW cancer; variant.  
XX OS Homo sapiens.  
XX FH Key Location/Qualifiers  
FT Misc-difference 685 /label= L685S  
FT /note= "wild-type Leu is replaced by Ser"  
FT Misc-difference 1282 /label= R1282A  
FT /note= "wild-type Arg is replaced by Ala"  
XX US6001563-A.  
XX 14-DEC-1999.  
XX 05-JUN-1995; 95US-00463179.  
XX 27-OCT-1992; 92US-00966923.  
XX 08-MAR-1993; 93US-00029340.  
XX 26-OCT-1993; 93US-00141893.  
XX 20-MAR-1995; 95US-00407207.  
XX (TOOH ) UNIV QUEENS KINGSTON.  
XX Cole SP, Deeley RG;  
XX WPI; 2000-061877/05.  
XX N-PSDB; AA239556.  
XX Identification of chemosensitizers useful for treating cancer, using  
XX nucleic acids encoding multidrug resistance-associated protein.  
XX Claim 3; Col 69-80; 77pp; English.  
XX The invention provides a method for identifying a substance which is a  
XX chemosensitizer that comprises, contacting a cell transfected with  
XX nucleic acid encoding multidrug resistance-associated protein (MRP) with  
XX a therapeutic agent in vitro. The method is useful for identifying  
XX chemosensitizers which may then be used to treat cancer (especially lung  
XX cancer). The method allows the identification of chemosensitizers which  
XX do not reverse P-glycoprotein-mediated multidrug resistance. The present  
XX sequence represents a human MRP variant  
XX Sequence 1531 AA;  
SQ

Query Match 79.4%; Score 7860; DB 3; Length 1531;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MALRGFCSDGSDPLNDWNVNTWNSPDFTKCFQNTLVWVPCFYLWACFPFFYLYLSRH 60  
Db 1 MALRGFCSDGSDPLNDWNVNTWNSPDFTKCFQNTLVWVPCFYLWACFPFFYLYLSRH 60  
Qy 61 DRGIQMTPLNKTALGFLWVWADLFYSFWSRSRGI FLAPVFLVSPILLGTTLLA 120  
Db 61 DRGIQMTPLNKTALGFLWVWADLFYSFWSRSRGI FLAPVFLVSPILLGTTLLA 120  
Qy 121 TFLQLERRKGVQSSGIMLTFWLVALVCAALILRSKIMTALKEDAQVDLFRDITFVYVS 180  
Db 121 TFLQLERRKGVQSSGIMLTFWLVALVCAALILRSKIMTALKEDAQVDLFRDITFVYVS 180  
Qy 181 LLLIQLVLSCFSDRSPLESETHDNPSPESASFLSRIITFWITGLIVRGYRQLEGSD 240  
Db 181 LLLIQLVLSCFSDRSPLESETHDNPSPESASFLSRIITFWITGLIVRGYRQLEGSD 240  
Qy 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKVYSSKDPAPQKESKVDANEVEAL 300  
Db 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKVYSSKDPAPQKESKVDANEVEAL 300  
Qy 301 IVKSPQKWNPSLFKVLKYKTFGPYFLMSFFPKAIHDLMMFSGPQILKLLIKPVNDTKAPD 360  
Db 301 IVKSPQKWNPSLFKVLKYKTFGPYFLMSFFPKAIHDLMMFSGPQILKLLIKPVNDTKAPD 360  
Qy 361 WQGYFTVLLFVTAQLQTLVHLQYFHCVSGMRIKTAVIGAVYRKALVITNSAKSSTV 420  
Db 361 WQGYFTVLLFVTAQLQTLVHLQYFHCVSGMRIKTAVIGAVYRKALVITNSAKSSTV 420  
Qy 421 GEIVNLSVDAQRFMDLATYINMIWSAPLQVILALYLLNLGSPVLAGVAVMLMVPVN 480  
Db 421 GEIVNLSVDAQRFMDLATYINMIWSAPLQVILALYLLNLGSPVLAGVAVMLMVPVN 480  
Qy 481 AVMAKTKTYQVAHMKSKDNRI KLMNEILNGIKVLKYAWELAFKDKVLAIROBELKVLK 540  
Db 481 AVMAKTKTYQVAHMKSKDNRI KLMNEILNGIKVLKYAWELAFKDKVLAIROBELKVLK 540  
Qy 541 KSAYLSAVGTTTWTCTPFLVALCTFAVVTVDENNILDAQTAFVLSALFNILRFPNLILP 600  
Db 541 KSAYLSAVGTTTWTCTPFLVALCTFAVVTVDENNILDAQTAFVLSALFNILRFPNLILP 600  
Qy 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWASDPPT 660  
Db 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWASDPPT 660  
Qy 661 LNGITFSIPEGALVAVVGQVCGKSSLLSALLAEMDKVEGHVAIKGSVAYVPPQAWIQND 720  
Db 661 LNGITFSIPEGALVAVVGQVCGKSSLLSALLAEMDKVEGHVAIKGSVAYVPPQAWIQND 720  
Qy 721 SIRENIFGCOLLEPYRSVIOACALLPDLLEILPSGDRTEIGEGKGNISGGOKQVSLAR 780  
Db 721 SIRENIFGCOLLEPYRSVIOACALLPDLLEILPSGDRTEIGEGKGNISGGOKQVSLAR 780  
Qy 781 AVYSNADILYLPDDPLSADVAHVGHKIFENVIGPKGMLKNKTRILVTHSHSYLPQVDVIV 840  
Db 781 AVYSNADILYLPDDPLSADVAHVGHKIFENVIGPKGMLKNKTRILVTHSHSYLPQVDVIV 840  
Qy 841 MSGGKISMGYSQELLARDGAPAEFLRTYASTEQEQDABENGVTGVSGPKEAKQENGM 900  
Db 841 MSGGKISMGYSQELLARDGAPAEFLRTYASTEQEQDABENGVTGVSGPKEAKQENGM 900  
Qy 901 LVTDAGLQRLQSSSSSYSGDISRHNSIAELOKAKKEETWKLMEADKAQTGVKL 960  
Db 901 LVTDAGLQRLQSSSSSYSGDISRHNSIAELOKAKKEETWKLMEADKAQTGVKL 960  
Qy 961 SVYDYMKAIGLFIISIFLIFMCMNVHSALASNYLSLWTDPIVNGTQHTKVLRSVYG 1020  
Db 961 SVYDYMKAIGLFIISIFLIFMCMNVHSALASNYLSLWTDPIVNGTQHTKVLRSVYG 1020  
Qy 1021 ALGISQGIAGVFGYSMAVSIIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRSKEL 1080

Db 1021 ALGISQIAVFGYSMAVSGIGILASRCLHVDLHLSILRSPMSFFERTPSGNLVNRFSKEL 1080  
 Qy 1081 DTVDMSIPVIMKFMGSLFNWIGACIVILLATPIAAIIPIPLGLIYFFVQRPYVASSRQL 1140  
 Db 1081 DTVDMSIPVIMKFMGSLFNWIGACIVILLATPIAAIIPIPLGLIYFFVQRPYVASSRQL 1140  
 Qy 1141 KELESVSRSPVYSHENETLGVSVIRAEQEERFHQSCLKVDENKQAYPSIVANRWLA 1200  
 Db 1141 KELESVSRSPVYSHENETLGVSVIRAEQEERFHQSCLKVDENKQAYPSIVANRWLA 1200  
 Qy 1201 VRLECVGNCIVLFAALFAVISRHSLSAGLVGSYSLSQVTTTLNWLVRMSSEMETNIVA 1260  
 Db 1201 VRLECVGNCIVLFAALFAVISRHSLSAGLVGSYSLSQVTTTLNWLVRMSSEMETNIVA 1260  
 Qy 1261 VERLKEYSETKEAPWQIQTAPPSSWPQVGRVFNRYCLRVREDLDFVLRHINVTINGG 1320  
 Db 1261 VERLKEYSETKEAPWQIQTAPPSSWPQVGRVFNRYCLRVREDLDFVLRHINVTINGG 1320  
 Qy 1321 EKVGIIVGRTGACKSLITLGLPRINSAGEIIIDGINIAKIGLHDLRFKTIIPDDPVLF 1380  
 Db 1321 EKVGIIVGRTGACKSLITLGLPRINSAGEIIIDGINIAKIGLHDLRFKTIIPDDPVLF 1380  
 Qy 1381 SGSLRMNLDPFQSYSDSEEWTSLELAHLKDFVSALPDKLDHCBAGGENLSVGQRLVCL 1440  
 Db 1381 SGSLRMNLDPFQSYSDSEEWTSLELAHLKDFVSALPDKLDHCBAGGENLSVGQRLVCL 1440  
 Qy 1441 ABALLRKTKILVLDATAAVDLETDLLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 Db 1441 ABALLRKTKILVLDATAAVDLETDLLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 Qy 1501 DKGEIQEYGAQPSDLLQQRGLFYSMAKDAGLV 1531  
 Db 1501 DKGEIQEYGAQPSDLLQQRGLFYSMAKDAGLV 1531

RESULT 5

AAV78873  
 ID AAV78873 standard; protein; 1531 AA.  
 XX AAV78873;  
 AC AAV78873;  
 DT 19-MAY-2000 (first entry)  
 XX Multidrug resistance protein (MRP) natural variant amino acid sequence.  
 DE Multidrug resistance protein (MRP) natural variant amino acid sequence.  
 KW Multidrug resistance protein; MRP; human; anthracycline; Vinca alkaloid;  
 XW epipodophyllotoxin; cancer; leukaemia.  
 XX Homo sapiens.  
 OS US6025473-A.  
 PN 15-FEB-2000.  
 XX 05-JUN-1995; 95US-00461384.  
 XX 27-OCT-1992; 92US-00966923.  
 PR 08-MAR-1993; 93US-00029340.  
 PR 26-OCT-1993; 93US-00141893.  
 PR 20-MAR-1995; 95US-00407207.  
 XX (TOOH ) UNIV QUEBENS KINGSTON.  
 PA  
 XX Cole SPC, Deeley RG;  
 PI WPI; 2000-181838/16.  
 DR N-PSDB; AA290193.  
 XX Isolated protein conferring multidrug resistance, to at least two drugs  
 PT selected from anthracyclines, epipodophyllotoxins and Vinca alkaloids, on  
 PT a drug sensitive mammalian cell.  
 XX

PS Claim 10; Col 79-88; 78pp; English.  
 XX This sequence represents a human multidrug resistance protein (MRP)  
 CC natural variant amino acid sequence. The human MRP confers multidrug  
 CC resistance, including resistance to at least two drugs selected from  
 CC anthracyclines, epipodophyllotoxins and Vinca alkaloids, on a drug  
 CC sensitive mammalian cell, when the protein is expressed in the cell. The  
 CC multidrug resistance is not substantially reversed by chemosensitizers  
 CC which reverse P-glycoprotein-mediated multidrug resistance. The MRP  
 CC protein sequence can be used to generate antibodies against MRP. The MRP  
 CC protein and nucleotide sequences can be used in compositions which are  
 CC used to treat patients with tumours displaying multidrug resistance. The  
 CC compositions and methods of the invention can be used particularly to  
 CC treat breast cancer, leukaemias, fibrosarcomas, cervical cancer, gliomas,  
 CC thymomas, neuroblastomas, and lung cancer. Antibodies directed against  
 CC MRP can be used to inhibit the multidrug resistance of a multidrug  
 CC resistant cell  
 XX  
 SQ Sequence 1531 AA;  
 Query Match 79.4%; Score 7860; DB 3; Length 1531;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 MALRGFCSADGSDPLWDNVTWNTSNPDFTKCFQNTVLVWVPCFYLWACFPFFLYLSRH 60  
 Db 1 MALRGFCSADGSDPLWDNVTWNTSNPDFTKCFQNTVLVWVPCFYLWACFPFFLYLSRH 60  
 Qy 61 DRGYQMTPLNKTALGFLWLWVWADLFPYSFWSRGIFLAPVFLVSPITLLGTTLLA 120  
 Db 61 DRGYQMTPLNKTALGFLWLWVWADLFPYSFWSRGIFLAPVFLVSPITLLGTTLLA 120  
 Qy 121 TFLIQLERRKGVSQSGIMLTFWLVALCALAILRSKIMTALKEDAQVDLPDITFYVYFS 180  
 Db 121 TFLIQLERRKGVSQSGIMLTFWLVALCALAILRSKIMTALKEDAQVDLPDITFYVYFS 180  
 Qy 181 LLLIQLVLSLCPDRSPFSETHDPNCPSSASFLSRITFWITGLIVRGYRPLEGSD 240  
 Db 181 LLLIQLVLSLCPDRSPFSETHDPNCPSSASFLSRITFWITGLIVRGYRPLEGSD 240  
 Qy 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKVYSSKDDPAQPKSSKVDANEVEAL 300  
 Db 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKVYSSKDDPAQPKSSKVDANEVEAL 300  
 Qy 301 IVKSPQKEWNPSPKVLVKTGPGYFLMSFFPKAIDHLMFSGPOLIKLLIKFVNDTKAPD 360  
 Db 301 IVKSPQKEWNPSPKVLVKTGPGYFLMSFFPKAIDHLMFSGPOLIKLLIKFVNDTKAPD 360  
 Qy 361 WQGYFTVLLFVTACLTQLVHLQYFHICFVSGMRIKTAVIGAVYRKALVITNSARKSSTV 420  
 Db 361 WQGYFTVLLFVTACLTQLVHLQYFHICFVSGMRIKTAVIGAVYRKALVITNSARKSSTV 420  
 Qy 421 GEIVNLSVDAQRFMDLATYINMIWSAPLQVILALYLLNWLNGPSVLGAVVWMLVMPVN 480  
 Db 421 GEIVNLSVDAQRFMDLATYINMIWSAPLQVILALYLLNWLNGPSVLGAVVWMLVMPVN 480  
 Qy 481 AVMAWTKTYQVAHMKSKDNRIKLWNETLNGIKVLKLYANELAFKDKVLAIROBELKVLK 540  
 Db 481 AVMAWTKTYQVAHMKSKDNRIKLWNETLNGIKVLKLYANELAFKDKVLAIROBELKVLK 540  
 Qy 541 KSAYLSAVGTFTWVCTPFLVALCTFAVVYVITDNNILDAQTAFAVLSALFNILRPLNLP 600  
 Db 541 KSAYLSAVGTFTWVCTPFLVALCTFAVVYVITDNNILDAQTAFAVLSALFNILRPLNLP 600  
 Qy 601 MVISSIVQASVSLKRLRIFLSHEELEPDSITERRPVKGGGTNSITVRNATTTWARSDDPT 660  
 Db 601 MVISSIVQASVSLKRLRIFLSHEELEPDSITERRPVKGGGTNSITVRNATTTWARSDDPT 660  
 Qy 661 LNGTHFTPEGALVAVVGVCCKGSLLSALLAEMDKVEGHVATKGSVAVYVQQAQWIND 720  
 Db 661 LNGTHFTPEGALVAVVGVCCKGSLLSALLAEMDKVEGHVATKGSVAVYVQQAQWIND 720  
 Qy 721 SLRENILFGCOLLEFPYRSVIQACALLPDLILPSPDRTEIGKGVNLGGQKQKRVSLAR 780

Db 721 SIRENIFGCOLEBYYRVSIOACALLPDLPLPSGDRTEIGEGKGNLSGGQKQVSLAR 780  
Qy 781 AVYSNADIYLPDPLSANDVAVGKH FENVIGPKMLKNKTRILVTHSMSYLPQVDVLIIV 840  
Db 781 AVYSNADIYLPDPLSANDVAVGKH FENVIGPKMLKNKTRILVTHSMSYLPQVDVLIIV 840  
Qy 841 MSGKISMGSGYQELLARDGAFABFLRYASTEQDABEENGVTGVSFGPKGKAKOMNGM 900  
Db 841 MSGKISMGSGYQELLARDGAFABFLRYASTEQDABEENGVTGVSFGPKGKAKOMNGM 900  
Qy 901 LVTSAGKOLQRLSSSSYSYSDISRHNSTAELOKAEAKKEETWKLMEADKAQTGVKL 960  
Db 901 LVTSAGKOLQRLSSSSYSYSDISRHNSTAELOKAEAKKEETWKLMEADKAQTGVKL 960  
Qy 961 SVYDYMKAIGLFIISLFIPLCMNHVSALASNYLSLWTDPIVNGTQEHKVKRLSVYG 1020  
Db 961 SVYDYMKAIGLFIISLFIPLCMNHVSALASNYLSLWTDPIVNGTQEHKVKRLSVYG 1020  
Qy 1021 ALGISQGIAGVSGYMAVIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFKSEL 1080  
Db 1021 ALGISQGIAGVSGYMAVIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFKSEL 1080  
Qy 1081 DTVDNMIEVIMFGSLFNIVGACIVILLATPIAAIIIPPLGLIYFFVQRPYVASSRQL 1140  
Db 1081 DTVDNMIEVIMFGSLFNIVGACIVILLATPIAAIIIPPLGLIYFFVQRPYVASSRQL 1140  
Qy 1141 KRLESVSRPVYSHNETLLGVSVTRAEEQERFIHQSDLVKDNOKAYYPSIVANRWLA 1200  
Db 1141 KRLESVSRPVYSHNETLLGVSVTRAEEQERFIHQSDLVKDNOKAYYPSIVANRWLA 1200  
Qy 1201 VRLECVGNCIVLFAALFAVISRHSLSAGLVLSVSYSLQVTTYLNLVMSSEMETNIVA 1260  
Db 1201 VRLECVGNCIVLFAALFAVISRHSLSAGLVLSVSYSLQVTTYLNLVMSSEMETNIVA 1260  
Qy 1261 VERLKEYSETEKAPWQIQTAPPSPQVGVKRPVFNRYCLRYREDLDFVLRHINVTINGG 1320  
Db 1261 VERLKEYSETEKAPWQIQTAPPSPQVGVKRPVFNRYCLRYREDLDFVLRHINVTINGG 1320  
Qy 1321 EKVGIIVGRTGAGKSLTLGLFRINESAGEIIIDGINIAKIGLHDLRFRKTIIPQDPVLF 1380  
Db 1321 EKVGIIVGRTGAGKSLTLGLFRINESAGEIIIDGINIAKIGLHDLRFRKTIIPQDPVLF 1380  
Qy 1381 SGLSRMNLDPFQSYDEEVTWTSLELAHLKDFVSALPDKLDHCEAGGENLSVGQRLVCL 1440  
Db 1381 SGLSRMNLDPFQSYDEEVTWTSLELAHLKDFVSALPDKLDHCEAGGENLSVGQRLVCL 1440  
Qy 1441 APALLRKTKILVLDATAAVDLETDLIQSTIRTQFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
Db 1441 APALLRKTKILVLDATAAVDLETDLIQSTIRTQFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
Qy 1501 DKGEIOEYCAPSDLAQORGLFYSMAKDAGLV 1531  
Db 1501 DKGEIOEYCAPSDLAQORGLFYSMAKDAGLV 1531

RESULT 6

ABG61810  
ID ABG61810 standard; protein; 1531 AA.

XX AC ABG61810;

XX DT 15-AUG-2002 (first entry)

XX DE Prostate cancer-associated protein #11.

XX KW Prostate cancer; prostate tumour tissue; human; mammal; cytostatic.

XX OS Mammalia.

XX FN WO200230268-A2.

XX PD 18-APR-2002.

XX 12-OCT-2001; 2001WO-US032045.  
XX 13-OCT-2000; 2000US-00687576.  
PR 08-DEC-2000; 2000US-00733288.  
PR 08-DEC-2000; 2000US-00733742.  
PR 24-JAN-2001; 2001US-0263957P.  
PR 16-MAR-2001; 2001US-0276791P.  
PR 16-MAR-2001; 2001US-0276888P.  
PR 06-APR-2001; 2001US-0281922P.  
PR 24-APR-2001; 2001US-0286214P.  
PR 30-APR-2001; 2001US-00847046.  
PR 04-MAY-2001; 2001US-0288589P.  
XX (EOSB-) EOS BIOTECHNOLOGY INC.  
XX Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;  
XX WPI; 2002-471335/50.  
XX N-PSDB; ABK92125.  
DR Detecting a prostate cancer-associated transcript in a cell in a patient,  
XX useful for diagnosing prostate cancer (PC) or screening modulators of PC,  
XX by determining if prostate cancer-associated genes are expressed in a  
XX prostate tissue.  
XX Claim 27; Page 309; 436pp; English.  
XX The present invention relates to methods of detecting a prostate cancer-  
XX associated transcript in a cell from a patient. The method comprises  
XX contacting a biological sample from the patient with prostate cancer-  
XX associated polynucleotides (designated PC genes) that selectively  
XX hybridise to a sequence that is at least 80% identical to them. The  
XX prostate cancer-associated polynucleotide sequences are differentially  
XX expressed in prostate tumour tissue or in prostate cancer and are derived  
XX from the tissues of various organisms such as humans or other mammals  
XX (e.g. mice, sheep and dogs). The methods of the invention are useful for  
XX diagnosing and treating prostate cancer in mammals. The prostate cancer-  
XX associated genes are useful for diagnosing or treating prostate cancer,  
XX as well as for identifying modulators of prostate cancer or agents that  
XX inhibit prostate cancer. The nucleic acid sequences are particularly  
XX useful in gene therapy, as a vaccine or in antisense applications.  
XX ABG61800-ABG61944 represent prostate cancer-associated proteins  
XX Sequence 1531 AA;

Query Match 79.4%; Score 7860; DB 5; Length 1531;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MALRGFCSADGSDPLMDNVTWNTSNPDFTKCFQNTVLVWVPCFYLWACFPFFYLYLSRH 60  
Db 1 MALRGFCSADGSDPLMDNVTWNTSNPDFTKCFQNTVLVWVPCFYLWACFPFFYLYLSRH 60  
Qy 61 DRGYIQTPLNKTALGFLWIVCWADLFYSFWRSGIFLAPVFLVSPDLLGTTLLA 120  
Db 61 DRGYIQTPLNKTALGFLWIVCWADLFYSFWRSGIFLAPVFLVSPDLLGTTLLA 120  
Qy 121 TFLQLERRKGVQSSGIMLTFWLVALCALAILSKINTALKEDAQVDFLFRDITVWVYFS 180  
Db 121 TFLQLERRKGVQSSGIMLTFWLVALCALAILSKINTALKEDAQVDFLFRDITVWVYFS 180  
Qy 181 LLLQLVLSGSDRSPFSETHDNPCESSASFLSRITTWITGLIVRGVQRLGSD 240  
Db 181 LLLQLVLSGSDRSPFSETHDNPCESSASFLSRITTWITGLIVRGVQRLGSD 240  
Qy 241 LWSLNKEDTSEQVVPVLVKNWKKCAKTRKQPKVYVSSKQPAQPKSSKVDANEVEAL 300  
Db 241 LWSLNKEDTSEQVVPVLVKNWKKCAKTRKQPKVYVSSKQPAQPKSSKVDANEVEAL 300  
Qy 301 IVKSPQKEWNPFLKVLKTYGPIFLMSFFPKAIHDLMMFSGPQILKLLIKFVNDTKAPD 360  
Db 301 IVKSPQKEWNPFLKVLKTYGPIFLMSFFPKAIHDLMMFSGPQILKLLIKFVNDTKAPD 360

QY 361 WQGYFTVLLFVTACIQTLVHXYPHICFVSGMRKTAIVIGAVYKALVITNSARKSSTV 420  
 Db 361 WQGYFTVLLFVTACIQTLVHXYPHICFVSGMRKTAIVIGAVYKALVITNSARKSSTV 420  
 QY 421 GEIVNLSVDAQRFMDLATYINIMISAPLOVTLALYLLWNLGSPVLAGVAVMLVMPVN 480  
 Db 421 GEIVNLSVDAQRFMDLATYINIMISAPLOVTLALYLLWNLGSPVLAGVAVMLVMPVN 480  
 QY 481 AVAMKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAWELAFKDKVLAIROBELKVLK 540  
 Db 481 AVAMKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAWELAFKDKVLAIROBELKVLK 540  
 QY 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTDENNILDAQTAFAFSLALFNILRPLNLTP 600  
 Db 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTDENNILDAQTAFAFSLALFNILRPLNLTP 600  
 QY 601 MVISSIVQASVSLKELRIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWARSDDPT 660  
 Db 601 MVISSIVQASVSLKELRIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWARSDDPT 660  
 QY 661 LINGITFSIPEGALVAVGVQGVGKSSLLSALLAEMDKVEGHVAIKGSVAVYVPPQAWIQND 720  
 Db 661 LINGITFSIPEGALVAVGVQGVGKSSLLSALLAEMDKVEGHVAIKGSVAVYVPPQAWIQND 720  
 QY 721 SLRENILFGCOLLEPYYSVIOACALLPDLETLPSGDRTEICEKGNVLSGGQKQVSLAR 780  
 Db 721 SLRENILFGCOLLEPYYSVIOACALLPDLETLPSGDRTEICEKGNVLSGGQKQVSLAR 780  
 QY 781 AVYSNADIYLFDDPLSADVAHVGHKIFENVIGPKGMLKXKTRILVTHSMSYLPQVDVIIV 840  
 Db 781 AVYSNADIYLFDDPLSADVAHVGHKIFENVIGPKGMLKXKTRILVTHSMSYLPQVDVIIV 840  
 QY 841 MSGGKISEMGYSQELLARDGAFAPFLRTYASTEQDAEENGVTGVSQPGKEAKQMGNGM 900  
 Db 841 MSGGKISEMGYSQELLARDGAFAPFLRTYASTEQDAEENGVTGVSQPGKEAKQMGNGM 900  
 QY 901 LVTDSAGLQORQLSSSSYSYSGDISRHNSTAELOKAEKEETWKLMEADKAOTGOVKL 960  
 Db 901 LVTDSAGLQORQLSSSSYSYSGDISRHNSTAELOKAEKEETWKLMEADKAOTGOVKL 960  
 QY 961 SVYDYMKAIGLFTSFLIFLMCNHVSALASNYWLSLWTDPIVNGTQEHTKVRLSVYG 1020  
 Db 961 SVYDYMKAIGLFTSFLIFLMCNHVSALASNYWLSLWTDPIVNGTQEHTKVRLSVYG 1020  
 QY 1021 ALGISQGIATVFGYSNAVSGIGILASRCLHVDLLHSILRSPMSFFERTPSGNLVRFSKEL 1080  
 Db 1021 ALGISQGIATVFGYSNAVSGIGILASRCLHVDLLHSILRSPMSFFERTPSGNLVRFSKEL 1080  
 QY 1081 DTVDMSIPEVIMKFMGSLFNVIGACIVILLATPTAAIIIPPLGLIYFFVQRFYVASSRQL 1140  
 Db 1081 DTVDMSIPEVIMKFMGSLFNVIGACIVILLATPTAAIIIPPLGLIYFFVQRFYVASSRQL 1140  
 QY 1141 KRLESVSRSPVYSHFNETHLLGVSVIRAFEEQERFIHQSDLKVDENQKAYYPSIVANRWLA 1200  
 Db 1141 KRLESVSRSPVYSHFNETHLLGVSVIRAFEEQERFIHQSDLKVDENQKAYYPSIVANRWLA 1200  
 QY 1201 VRLECVGNCIVLFAALPAVIRSRHLSAGLVLSYSLQVTTYLNLVWRMSSEMETNIVA 1260  
 Db 1201 VRLECVGNCIVLFAALPAVIRSRHLSAGLVLSYSLQVTTYLNLVWRMSSEMETNIVA 1260  
 QY 1261 VERLKEYSETEKAPWQIQETAPSSWPQVGRVFRNCLRYREDLDFVLRIHNTVINGG 1320  
 Db 1261 VERLKEYSETEKAPWQIQETAPSSWPQVGRVFRNCLRYREDLDFVLRIHNTVINGG 1320  
 QY 1321 EKVGVIGRTGAGKSSLTGLFRINESARGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
 Db 1321 EKVGVIGRTGAGKSSLTGLFRINESARGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
 QY 1381 SGSLRMNLDPQSQYSDDEWVTSLELAHLKDFVSALPDKLDHECAGGENLSVGQRLVCL 1440  
 Db 1381 SGSLRMNLDPQSQYSDDEWVTSLELAHLKDFVSALPDKLDHECAGGENLSVGQRLVCL 1440

QY 1441 ARALLRKTILVLDEATAAVDLETTDLIQSTIRIQFEDCTVLTIAHRLNTIMDYTRVILV 1500  
 Db 1441 ARALLRKTILVLDEATAAVDLETTDLIQSTIRIQFEDCTVLTIAHRLNTIMDYTRVILV 1500  
 QY 1501 DKGBIQEYGAPESDLLQQRGLFYSMKADAGLV 1531  
 Db 1501 DKGBIQEYGAPESDLLQQRGLFYSMKADAGLV 1531

RESULT 7  
 ABM35012  
 ID ABM35012 standard; protein; 1531 AA.  
 XX  
 AC ABM35012;  
 XX  
 DT 08-OCT-2003 (first entry)  
 XX  
 DE Cancer based on CYP3A5 related protein SEQ ID NO:678.  
 XX  
 KW Cancer; CYP3A5; irinotecan; pharmaceutical; malignant glioma;  
 KW cytochrome p450; subfamily IIIA; nifedipine oxidase; polypeptide 5;  
 KW cytosstatic.  
 XX  
 OS Unidentified.  
 XX  
 PN WO2003013534-A2.  
 XX  
 PD 20-FEB-2003.  
 XX  
 PF 23-JUL-2002; 2002WO-EP008219.  
 XX  
 PR 23-JUL-2001; 2001EP-00117608.  
 PR 24-MAY-2002; 2002EP-00011710.  
 XX  
 PA (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.  
 XX  
 PI Heinrich G, Kerb R;  
 XX  
 DR WPI; 2003-268144/26.  
 XX  
 PT New use of irinotecan for preparation of compositions for treating cancer  
 PT in subject having genome with variant allele comprising cytochrome p450,  
 PT subfamily IIIA, polypeptide 5 polynucleotide, termed CYP3A5.  
 XX  
 PS Disclosure; SEQ ID NO 678; 86pp; English.  
 XX  
 CC The present invention describes the use of irinotecan (I) or its  
 CC derivative for the preparation of a pharmaceutical composition for  
 CC treating colorectal, cervical, gastric, lung, ovarian or pancreatic  
 CC cancer, or malignant glioma in a subject having a genome with a variant  
 CC allele which comprises a cytochrome p450, subfamily IIIA (nifedipine  
 CC oxidase), polypeptide 5 (CYP3A5) polynucleotide (II). (I) and (II) have  
 CC cytostatic activity. The therapeutic applications of (I) is improved,  
 CC since it is possible to individually treat a subject with an appropriate  
 CC dosage and/or an appropriate derivative of (I). Therefore, undesirable,  
 CC harmful or toxic effects are efficiently avoided. Unnecessary and  
 CC potentially harmful treatment of those subjects who do not respond to the  
 CC treatment with substances (nonresponders), as well as the development of  
 CC drug resistances due to suboptimal drug dosing can be avoided. AC62200  
 CC to AC62751 and ABM34912 to ABM35013 represent sequences used in the  
 CC exemplification of the present invention  
 XX  
 SQ Sequence 1531 AA;

Query Match 79.4%; Score 7860; DB 6; Length 1531;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCSADGSDPLMDWNTWNTSNDFTKCFONTVLVWVPCFYLWACPFYFLYSRH 60  
 Db 1 MALRGFCSADGSDPLMDWNTWNTSNDFTKCFONTVLVWVPCFYLWACPFYFLYSRH 60  
 QY 61 DRGYIQMTPLNKTALGFLLIWVCWADLFYSFWSRSGIFLAPVFLVSPLLGTTILLA 120

Db 61 DRGYIQMTPLNKTALGFLLLIVCWADLFYFWSRSGIFLAPVFLVSPTLTGITLLA 120  
Qy 121 TELIOLERRKGVOSSGIMLTFWLVALVCAALATLRSKIMTALKEDAQVDLFRDITFVYVS 180  
Db 121 TELIOLERRKGVOSSGIMLTFWLVALVCAALATLRSKIMTALKEDAQVDLFRDITFVYVS 180  
Qy 181 LLLIQLVLSCFSDRSPFLSETIHDNPNCPSESSASFLSRITFWITGLIVRGYRQPLEGSD 240  
Db 181 LLLIQLVLSCFSDRSPFLSETIHDNPNCPSESSASFLSRITFWITGLIVRGYRQPLEGSD 240  
Qy 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKVVYSSKDPAPKESKVDANEVEAL 300  
Db 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKVVYSSKDPAPKESKVDANEVEAL 300  
Qy 301 IVKSPQKWNPSLFVLYKTFPGPYFLMSPFFKAIHDLMMFSGPOLKLLIKFVNDTKAPD 360  
Db 301 IVKSPQKWNPSLFVLYKTFPGPYFLMSPFFKAIHDLMMFSGPOLKLLIKFVNDTKAPD 360  
Qy 361 WQGYFYTVLLFVTAQLTLVLHQYFHCIFVSGMRKTAIGAVYRKALVITNSARKSSTV 420  
Db 361 WQGYFYTVLLFVTAQLTLVLHQYFHCIFVSGMRKTAIGAVYRKALVITNSARKSSTV 420  
Qy 421 GEIVNLSVDAQRFMDLATYINMWSAPLOVILALYLLNLGSPVLGAVVNLMPVN 480  
Db 421 GEIVNLSVDAQRFMDLATYINMWSAPLOVILALYLLNLGSPVLGAVVNLMPVN 480  
Qy 481 AVWAMKTKYQVAHWKSKDNRIKLWNEILNGIKVLKLYAWELA FKDKVLA ROELKVLK 540  
Db 481 AVWAMKTKYQVAHWKSKDNRIKLWNEILNGIKVLKLYAWELA FKDKVLA ROELKVLK 540  
Qy 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTDENNILDAQTAFAVSLALFNILRFPNLILP 600  
Db 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTDENNILDAQTAFAVSLALFNILRFPNLILP 600  
Qy 601 WYISSIVQASVSLKRLRIPLSHEELEPOSIERRRPVKGGTNSITVRNATFTWASDDPT 660  
Db 601 WYISSIVQASVSLKRLRIPLSHEELEPOSIERRRPVKGGTNSITVRNATFTWASDDPT 660  
Qy 661 LNGITFTSPEGALVAVGQVGCGKSLLSALLAEMDKVEGHVAIRKGSVAVYVPOQAWIOND 720  
Db 661 LNGITFTSPEGALVAVGQVGCGKSLLSALLAEMDKVEGHVAIRKGSVAVYVPOQAWIOND 720  
Qy 721 SLRENILFGCOLLEPYRYSVIOACALLPDLLEILPSGDRTEIGEKGVLNLSGGQKQVSLAR 780  
Db 721 SLRENILFGCOLLEPYRYSVIOACALLPDLLEILPSGDRTEIGEKGVLNLSGGQKQVSLAR 780  
Qy 781 AVYSNADIYLFDDPLSADVAHVGHKHFENVIGPKGMLKNKTRILVTHSMSYLPQVDVIV 840  
Db 781 AVYSNADIYLFDDPLSADVAHVGHKHFENVIGPKGMLKNKTRILVTHSMSYLPQVDVIV 840  
Qy 841 MSGGKISEGYSOELLARDGAFAELRTVASTEQEDAENGVTGSGPKEAKOMENG 900  
Db 841 MSGGKISEGYSOELLARDGAFAELRTVASTEQEDAENGVTGSGPKEAKOMENG 900  
Qy 901 LVTDSAGLQORQLSSSSYSYSDISRHNSTAELOKAEAKKEETWKLMEADKAQTQGVKL 960  
Db 901 LVTDSAGLQORQLSSSSYSYSDISRHNSTAELOKAEAKKEETWKLMEADKAQTQGVKL 960  
Qy 961 SVYDYMKAIGLIFSLFIFPMCNHVSALASNYLSLWTDPIVNGTQEHTKVRLSVYG 1020  
Db 961 SVYDYMKAIGLIFSLFIFPMCNHVSALASNYLSLWTDPIVNGTQEHTKVRLSVYG 1020  
Qy 1021 ALGISQGIATVFGYSMAVSTGGILASRCLHVDLLHSILRSPMSEFFERTSGNIVNPFSEL 1080  
Db 1021 ALGISQGIATVFGYSMAVSTGGILASRCLHVDLLHSILRSPMSEFFERTSGNIVNPFSEL 1080  
Qy 1081 DTVDMSIPEVIMKMGSLFNWIGACIVILLATPIAAIIPPLGLIYFFVQRYVASSRQL 1140  
Db 1081 DTVDMSIPEVIMKMGSLFNWIGACIVILLATPIAAIIPPLGLIYFFVQRYVASSRQL 1140  
Qy 1141 KRLESVRSRSPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENQKAYYPSIVANRWLA 1200

Db 1141 KRLESVRSRSPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENQKAYYPSIVANRWLA 1200  
Qy 1201 VRLECVGNICIVLFAALFAVISRHSLSAGLVLSVSYSQVTTYLNWLVMSSEMETNIVA 1260  
Db 1201 VRLECVGNICIVLFAALFAVISRHSLSAGLVLSVSYSQVTTYLNWLVMSSEMETNIVA 1260  
Qy 1261 VERLKEYSETKEAPWQIQETAPPSSWPQVGRVEFRNRYCLRYREDLDFVLRHINVTINGG 1320  
Db 1261 VERLKEYSETKEAPWQIQETAPPSSWPQVGRVEFRNRYCLRYREDLDFVLRHINVTINGG 1320  
Qy 1321 EKVGVGTGCTGAGKSSLTGLFRINESAGEEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
Db 1321 EKVGVGTGCTGAGKSSLTGLFRINESAGEEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
Qy 1381 SGLSRMNLDPFSQYSDEEVTWTSLELAHLKDFVSALPDKLDHECAGGENLSVGOQLVCL 1440  
Db 1381 SGLSRMNLDPFSQYSDEEVTWTSLELAHLKDFVSALPDKLDHECAGGENLSVGOQLVCL 1440  
Qy 1441 ARALLRKTKILVLEATAAVALDETDLLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
Db 1441 ARALLRKTKILVLEATAAVALDETDLLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
Qy 1501 DKGEIOEQYCAPSDLLIQORGLFYSMAKDAGLV 1531  
Db 1501 DKGEIOEQYCAPSDLLIQORGLFYSMAKDAGLV 1531

RESULT 8  
ADB20865

ID ADB20865 standard; protein; 1531 AA.

XX AC ADB20865;

XX DT 20-NOV-2003 (first entry)

XX DE MRPI based cancer related protein SEQ ID NO:678.

XX KW irinotecan; colorectal cancer; cervical cancer; gastric cancer;  
XX KW lung cancer; ovarian cancer; pancreatic cancer; malignant glioma;  
XX KW variant allele; multidrug resistance protein 1; MRPI; cytostatic.

XX OS Unidentified.

XX PN WO2003013533-A2.

XX PD 20-FEB-2003.

XX PF 23-JUL-2002; 2002WO-EP008200.

XX PR 23-JUL-2001; 2001EP-00117608.

XX PR 24-MAY-2002; 2002EP-00011710.

XX PA (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.

XX PI Heinrich G, Kerb R;

XX WPI; 2003-354397/33.

XX PT Use of irinotecan or its derivative for preparation of a pharmaceutical  
XX PT composition for treating cancer in a subject having a genome with a  
XX PT variant allele comprising a multidrug resistance protein 1  
XX PT polynucleotide.

XX PS Disclosure; SEQ ID NO 678; 100pp; English.

XX CC The present invention describes a method for the use of irinotecan (I) or  
XX CC its derivative for the preparation of a pharmaceutical composition for  
XX CC treating colorectal, cervical, gastric, lung, ovarian or pancreatic  
XX CC cancer, or malignant glioma in a subject having a genome with a variant  
XX CC allele which comprises a multidrug resistance protein 1 (MRPI)  
XX CC polynucleotide (II). (I) has cytostatic activity. (II) or its derivative  
XX CC can be used for the preparation of a pharmaceutical composition for  
XX CC treating colorectal, cervical, gastric, lung, ovarian or pancreatic

CC cancer, or malignant glioma in a subject, where the subject is a human  
CC (preferably African or Asian) or a mouse. The present sequence represents  
CC a sequence which is used in the exemplification of the present invention.  
XX  
SQ Sequence 1531 AA;

Query Match 79.4%; Score 7860; DB 6; Length 1531;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MALRFGCSADGSDPLWDMNVNTWNTSNPDTKCFQNTVLVWVPCFYLWACFPFYLYLSRH	60
DB	1	MALRFGCSADGSDPLWDMNVNTWNTSNPDTKCFQNTVLVWVPCFYLWACFPFYLYLSRH	60
QY	61	DRGYIQMTPLNKTALGFLMIWCVADLFYSFWSRSGIFLAPFLVPSLTLLGTTLLA	120
DB	61	DRGYIQMTPLNKTALGFLMIWCVADLFYSFWSRSGIFLAPFLVPSLTLLGTTLLA	120
QY	121	TFLIQLERRKGVOSSGIMLTFWLVALCALAIRSKIMTALKEDAQVDFRDITFYVYFS	180
DB	121	TFLIQLERRKGVOSSGIMLTFWLVALCALAIRSKIMTALKEDAQVDFRDITFYVYFS	180
QY	181	LLLIQLVLSFSDRSPLSETIHDNPNCPSSASFLSRITFWITGLIVRGVROPLEGSD	240
DB	181	LLLIQLVLSFSDRSPLSETIHDNPNCPSSASFLSRITFWITGLIVRGVROPLEGSD	240
QY	241	LWSLNKEDTSEOVVPLVKNWKECAKTRKOPKVYVSSKOPAKPKSSKVDANEEVEAL	300
DB	241	LWSLNKEDTSEOVVPLVKNWKECAKTRKOPKVYVSSKOPAKPKSSKVDANEEVEAL	300
QY	301	IVKSPQKWNPSLFKVLKTKGYPFLMSFFFKAIHDLMMFSGPQLLKLIKFNVDTKAPD	360
DB	301	IVKSPQKWNPSLFKVLKTKGYPFLMSFFFKAIHDLMMFSGPQLLKLIKFNVDTKAPD	360
QY	361	WQGYFYTVLLFTVACLOTLVHLYPHICFVSGMRITKAVIGAVRKALVITNSARKSTV	420
DB	361	WQGYFYTVLLFTVACLOTLVHLYPHICFVSGMRITKAVIGAVRKALVITNSARKSTV	420
QY	421	GEIVNLSVDAQRFMDLATYINMWSAPLOVILALYLNLNLPVSLAGVAVMLWVPVN	480
DB	421	GEIVNLSVDAQRFMDLATYINMWSAPLOVILALYLNLNLPVSLAGVAVMLWVPVN	480
QY	481	AVAMKTKTYQVAHMKSDNRKILKMEILNGIKVLGYAWELAFKDKVLAIROBELKVLK	540
DB	481	AVAMKTKTYQVAHMKSDNRKILKMEILNGIKVLGYAWELAFKDKVLAIROBELKVLK	540
QY	541	KSAYLSAVGTFTWCTPFLVALCTFAVYVTTDENNILDQAQTAFAVSLALFNILRPLNLP	600
DB	541	KSAYLSAVGTFTWCTPFLVALCTFAVYVTTDENNILDQAQTAFAVSLALFNILRPLNLP	600
QY	601	MVISIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWARSDDPT	660
DB	601	MVISIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWARSDDPT	660
QY	661	INGITFSIPEGALNAVVGQCGKSSLLSALLAEMDKVEGHVAIKGSVAYVPPQAWIQND	720
DB	661	INGITFSIPEGALNAVVGQCGKSSLLSALLAEMDKVEGHVAIKGSVAYVPPQAWIQND	720
QY	721	SIRENILEGCOLEBPYRSVTOACALLPDILPSGDRTEIGKGVNLSGQKQKVSLAR	780
DB	721	SIRENILEGCOLEBPYRSVTOACALLPDILPSGDRTEIGKGVNLSGQKQKVSLAR	780
QY	781	AVYSNADITYLFDPLSADVAHVGHIFENVIKPGKMLKNKTRILVTHSMSYLPQVDVLIIV	840
DB	781	AVYSNADITYLFDPLSADVAHVGHIFENVIKPGKMLKNKTRILVTHSMSYLPQVDVLIIV	840
QY	841	MSGGKISMSGYQELLARDGAFABFLRTYASTEQDABENGVTGVSGFGKEAKQOMENGM	900
DB	841	MSGGKISMSGYQELLARDGAFABFLRTYASTEQDABENGVTGVSGFGKEAKQOMENGM	900
QY	901	LVTDSAGKOLOROLSSSSSSSGDLSRHNSYAELOKABAKKEETWKLMEADKATQGVKL	960
DB	901	LVTDSAGKOLOROLSSSSSSSGDLSRHNSYAELOKABAKKEETWKLMEADKATQGVKL	960

QY	961	SVTYDMKAIAGLISFLSIFLFCMCHVHSALASNYWLSLWTDPIVNGTOEHTKVPLSVYG	1020
DB	961	SVTYDMKAIAGLISFLSIFLFCMCHVHSALASNYWLSLWTDPIVNGTOEHTKVPLSVYG	1020
QY	1021	ALGISQGIAGVFGYSMAVSIIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFSKEL	1080
DB	1021	ALGISQGIAGVFGYSMAVSIIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFSKEL	1080
QY	1081	DTVDSMIDPEVTKMPGSLFNFGVIGACIVILLATPTAAIIIPPLGLIYFFVQFVYVASSRQL	1140
DB	1081	DTVDSMIDPEVTKMPGSLFNFGVIGACIVILLATPTAAIIIPPLGLIYFFVQFVYVASSRQL	1140
QY	1141	KRLSVSRSPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENOKAYYPSIVANRWLA	1200
DB	1141	KRLSVSRSPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENOKAYYPSIVANRWLA	1200
QY	1201	VRLECVGNCIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTYLNWLVRMSSEMETNIVA	1260
DB	1201	VRLECVGNCIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTYLNWLVRMSSEMETNIVA	1260
QY	1261	VERLKEYSETEKEAPWQIQETAPPSSWPQVGRVFRNYCLRYREDLDPVLRHINVTINGG	1320
DB	1261	VERLKEYSETEKEAPWQIQETAPPSSWPQVGRVFRNYCLRYREDLDPVLRHINVTINGG	1320
QY	1321	EKVGIVGRTGAGKSSLTGLFRINESAEGETIIDGINIAKIGLHDLRFKTIIPQDPVLF	1380
DB	1321	EKVGIVGRTGAGKSSLTGLFRINESAEGETIIDGINIAKIGLHDLRFKTIIPQDPVLF	1380
QY	1381	SGSLRMNLDPFSQYSDSEVWTSLELAHLKDFVSALPKLDHECAEGGENLSVGORQLVCL	1440
DB	1381	SGSLRMNLDPFSQYSDSEVWTSLELAHLKDFVSALPKLDHECAEGGENLSVGORQLVCL	1440
QY	1441	ARALLRTKILVLDEATAAVDLETDLLIOSTIRTOFEDCTVLTIAHRLNTMDYTRVIVL	1500
DB	1441	ARALLRTKILVLDEATAAVDLETDLLIOSTIRTOFEDCTVLTIAHRLNTMDYTRVIVL	1500
QY	1501	DKGIIQEVGAPSDLLQORGLFYSMAXDAGLV	1531
DB	1501	DKGIIQEVGAPSDLLQORGLFYSMAXDAGLV	1531

RESULT 9

ADB87954

ID ADB87954 standard; protein; 1531 AA.

AC ADB87954;

XX  
DT 04-DEC-2003 (first entry)

XX Human UGT1A1 protein sequence SEQ ID NO:678.

DE  
XX  
KW irinotecan; cancer; UGT1A1; cytostatic; topoisomerase I inhibitor;  
KW colorectal cancer; cervical cancer; gastric cancer; lung cancer;  
KW ovarian cancer; pancreatic cancer; malignant glioma;  
KW uridine diphosphate glycosyltransferase I member A1.

OS Homo sapiens.

FN WO2003013536-A2.

XX 20-FEB-2003.

XX 23-JUL-2002; 2002WO-EP008217.

XX 23-JUL-2001; 2001EP-00117608.

PR 24-MAY-2002; 2002EP-00011710.

XX (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.

XX Heinrich G, Kerb R;

XX WPI; 2003-289896/28.

XX Use of irinotecan to treat cancer patient by determining if patient has  
PT variant alleles of UGT1A1 gene, administering increased/decreased amounts  
PT of irinotecan based on increased/decreased levels of UGT1A1 gene product.  
XX  
PS Disclosure; SEQ ID NO 678; 107pp; English.  
XX  
CC The invention relates to the novel use of irinotecan to treat a patient  
CC suffering from cancer. This involves determining if the patient has one  
CC or more variant alleles of the UGT1A1 gene, and if the patient has one or  
CC more of such variant alleles, irinotecan is administered in an increased  
CC or decreased amount in comparison to the amount that is administered  
CC without regard to the patient's alleles in the UGT1A1 gene. The invention  
CC has cytostatic activity. A composition of the invention acts as a  
CC topoisomerase I inhibitor. The method is useful for treating a patient,  
CC an animal e.g. mouse or a human, preferably African or Asian, suffering  
CC from cancer such as colorectal, cervical, gastric cancer, lung, ovarian,  
CC pancreatic cancer or malignant glioma. The present sequence is udes in  
CC the exemplification of the invention.  
XX  
SQ Sequence 1531 AA;

Query Match 79.4%; Score 7860; DB 7; Length 1531;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCGADGSDPLWDMNVTWNTSNPDFTKCFQNTVLVWVPCFYLWACFPFYLILSRH 60  
DB 1 MALRGFCGADGSDPLWDMNVTWNTSNPDFTKCFQNTVLVWVPCFYLWACFPFYLILSRH 60

QY 61 DRGYIQMTPLNKTALGFLWIVCWADLFYSEWERSRGIFLAPVFLVSPDLLGTTLLA 120  
DB 61 DRGYIQMTPLNKTALGFLWIVCWADLFYSEWERSRGIFLAPVFLVSPDLLGTTLLA 120

QY 121 TELIQLERRKGQSSGIMLTFWLVALVCMALILRSKIMTALKEDAQVDLFRDITVYVFS 180  
DB 121 TELIQLERRKGQSSGIMLTFWLVALVCMALILRSKIMTALKEDAQVDLFRDITVYVFS 180

QY 181 LLLIQLVLSFCSDRPLSFSETIHDNPPCPSSASFLSRTTFWITGLVGRVQPLEGSD 240  
DB 181 LLLIQLVLSFCSDRPLSFSETIHDNPPCPSSASFLSRTTFWITGLVGRVQPLEGSD 240

QY 241 LWSLNKEDTSEQVVPVLVKNMKECAKTRKQPVVYVYSSKDPAPKSSKVDANEVEAL 300  
DB 241 LWSLNKEDTSEQVVPVLVKNMKECAKTRKQPVVYVYSSKDPAPKSSKVDANEVEAL 300

QY 301 IVKSPQKEWNPFLFKVLYKTFPGYFLMSFFFKAIHDLMMFSGPQILKLLIKFVNDTKAPD 360  
DB 301 IVKSPQKEWNPFLFKVLYKTFPGYFLMSFFFKAIHDLMMFSGPQILKLLIKFVNDTKAPD 360

QY 361 WQGYFVTVLLFVTACIQTLVHGYPHICFVSGMRKTAIVGAVYKALVITNSARKSTV 420  
DB 361 WQGYFVTVLLFVTACIQTLVHGYPHICFVSGMRKTAIVGAVYKALVITNSARKSTV 420

QY 421 GEIVNLMSVDAQRFMDLATYINMWSAPLQVILALYLLWNLGSPVLAGVAVMVLMPVN 480  
DB 421 GEIVNLMSVDAQRFMDLATYINMWSAPLQVILALYLLWNLGSPVLAGVAVMVLMPVN 480

QY 481 AYAMKTKTYQVAHMKSKDNRIKLMEILNGIKVLKYAWELAPKDKVLAIRQEBELKVLK 540  
DB 481 AYAMKTKTYQVAHMKSKDNRIKLMEILNGIKVLKYAWELAPKDKVLAIRQEBELKVLK 540

QY 541 KSAYLSAVGTFTWCTTPPLVALCTFAVYVITDENNILDQAOTAFVSLALFNILRFPNLILP 600  
DB 541 KSAYLSAVGTFTWCTTPPLVALCTFAVYVITDENNILDQAOTAFVSLALFNILRFPNLILP 600

QY 601 MYISSIVQASVLSKRLRIFLSHEELEPDSIERPPVKDGGGTSITVRNATFTWASDDPT 660  
DB 601 MYISSIVQASVLSKRLRIFLSHEELEPDSIERPPVKDGGGTSITVRNATFTWASDDPT 660

QY 661 LNGITFTSIPEGALVAVVGQVCGKSLLSALLAEMDKVEGHVAIKGSVAVYVQQAQIWD 720  
DB 661 LNGITFTSIPEGALVAVVGQVCGKSLLSALLAEMDKVEGHVAIKGSVAVYVQQAQIWD 720

QY 721 SILRENILFGCOLEBPYRSVIOACALLPDLLEILPSGDRTEIGEGKGNLSGGQKQVSLAR 780  
DB 721 SILRENILFGCOLEBPYRSVIOACALLPDLLEILPSGDRTEIGEGKGNLSGGQKQVSLAR 780

QY 781 AVYNSADIYLLPDDPLSAVDHVGKHI FENVIGPKGMLKNKTRILVTHSMYLPQVDVILV 840  
DB 781 AVYNSADIYLLPDDPLSAVDHVGKHI FENVIGPKGMLKNKTRILVTHSMYLPQVDVILV 840

QY 841 MSGGKISMGYSQELLARDGAPAEFLRYASTEQDQAEENGVTGSGPGKAKQWENGM 900  
DB 841 MSGGKISMGYSQELLARDGAPAEFLRYASTEQDQAEENGVTGSGPGKAKQWENGM 900

QY 901 LVTDSAGLQORQLSSSSSSYSGDISRHHNSTAEQLQKAEKBEETWKLMEADKAQTCQVKL 960  
DB 901 LVTDSAGLQORQLSSSSSSYSGDISRHHNSTAEQLQKAEKBEETWKLMEADKAQTCQVKL 960

QY 961 SVYWDYKKAIGLFTSIFLFCMCHVUSALASNYWLSLWTDPIVNGTQEHTKVLSVYG 1020  
DB 961 SVYWDYKKAIGLFTSIFLFCMCHVUSALASNYWLSLWTDPIVNGTQEHTKVLSVYG 1020

QY 1021 ALGISQGIATVFGYSMAVSIIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFSEL 1080  
DB 1021 ALGISQGIATVFGYSMAVSIIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFSEL 1080

QY 1081 DTVDSMIPEVIKMFNGSLFNVIGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQL 1140  
DB 1081 DTVDSMIPEVIKMFNGSLFNVIGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQL 1140

QY 1141 KRLSVSRSPVYSHNETLLGVSVIRAFEBQERFIHQSDLKVDENQKAYYPSIVANRWLA 1200  
DB 1141 KRLSVSRSPVYSHNETLLGVSVIRAFEBQERFIHQSDLKVDENQKAYYPSIVANRWLA 1200

QY 1201 VRLECVGNICVLPFAALPAVISRHSLSAGLVLSVSYSLQVTTYLNWLVMSSEMETNI VA 1260  
DB 1201 VRLECVGNICVLPFAALPAVISRHSLSAGLVLSVSYSLQVTTYLNWLVMSSEMETNI VA 1260

QY 1261 VERLKEYSETBEAPWQIQETAPPSWPQVGRVEFRNYCLARYREDLDFVLRHINVTINGG 1320  
DB 1261 VERLKEYSETBEAPWQIQETAPPSWPQVGRVEFRNYCLARYREDLDFVLRHINVTINGG 1320

QY 1321 EKVGIIVGTGAGKSSLTGLFRINESARBEIIIDGINIAKIGLHDLRFRKTIIPDPVLF 1380  
DB 1321 EKVGIIVGTGAGKSSLTGLFRINESARBEIIIDGINIAKIGLHDLRFRKTIIPDPVLF 1380

QY 1381 SGSLSRMNLDPFQSYSDSEEVWTSLELAHLKDFVSALPDKLDHECAGGENLSVQRLVCL 1440  
DB 1381 SGSLSRMNLDPFQSYSDSEEVWTSLELAHLKDFVSALPDKLDHECAGGENLSVQRLVCL 1440

QY 1441 ARALLRKTILVLDDEATAVLDLTDLLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIL 1500  
DB 1441 ARALLRKTILVLDDEATAVLDLTDLLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIL 1500

QY 1501 DKGEIQEYGAPESDLLQORGLFYMAKDAGLV 1531  
DB 1501 DKGEIQEYGAPESDLLQORGLFYMAKDAGLV 1531

RESULT 10  
ADB96937  
ID ADB96937 standard; protein; 1531 AA.  
XX  
AC ADB96937;  
XX  
DT 04-DEC-2003 (first entry)  
XX  
DE Human MDR1 related protein sequence SEQ ID NO:678.  
XX  
KW irinotecan; colorectal cancer; cervical cancer; gastric cancer;  
KW lung cancer; ovarian cancer; pancreatic cancer; malignant glioma;  
KW multidrug resistance 1; MDR1; cytostatic; human; CYP3A5; MRP1; MDR1;  
XX TOPI.

OS Homo sapiens.  
 PN WO2003013537-A2.  
 XX 20-FEB-2003.  
 XX 23-JUL-2002; 2002WO-BF008218.  
 XX 23-JUL-2001; 2001BP-00117608.  
 PR 24-MAY-2002; 2002EP-00011710.  
 XX (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.  
 XX Heinrich G, Kerb R;  
 XX WPI; 2003-268145/26.  
 XX  
 PT New use of irinotecan for preparation of pharmaceutical compositions for  
 PT treating cancer in subject having genome with variant allele comprising  
 PT multidrug resistance 1 polynucleotide.  
 XX Disclosure; SEQ ID NO 678; 130pp; English.  
 XX  
 CC The invention relates to the novel use of irinotecan or its derivative  
 CC for the preparation of pharmaceutical compositions for treating  
 CC colorectal, gastric, lung, ovarian or pancreatic cancer, or  
 CC malignant glioma in a subject having a genome with a variant allele which  
 CC comprises a multidrug resistance 1 (MDR1) polynucleotide. A composition  
 CC of the invention has cytostatic activity. The invention is useful for the  
 CC preparation of pharmaceutical compositions for treating colorectal,  
 CC cervical, gastric, lung, ovarian or pancreatic cancer, or malignant  
 CC glioma in a subject (preferably human, more preferably African or Asian)  
 CC or a mouse. The present sequence is used in the exemplification of the  
 CC invention.  
 XX  
 SQ Sequence 1531 AA;  
 Query Match 79.4%; Score 7860; DB 7; Length 1531;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 1 MALRGCADGSDPLDWNVTWNTSNPDTKCFQNTLVWVPCFVLMACFPFFYLYLSRH 60  
 1 MALRGCADGSDPLDWNVTWNTSNPDTKCFQNTLVWVPCFVLMACFPFFYLYLSRH 60  
 61 DRGYQMTPLNKTALGFLWVWADLFYSFWSRSGIFLAPVFLVSPITLGLITLLA 120  
 61 DRGYQMTPLNKTALGFLWVWADLFYSFWSRSGIFLAPVFLVSPITLGLITLLA 120  
 121 TFLIQLERRKGVQSSGIMLTFWLVALVCAALILRSKIMTALKEDAQVDLFRDITFYVYFS 180  
 121 TFLIQLERRKGVQSSGIMLTFWLVALVCAALILRSKIMTALKEDAQVDLFRDITFYVYFS 180  
 181 LLLIQLVLSFSDRSPLSETHDNPCESSASFLSRITFWITGLIVRGYRPLEGSD 240  
 181 LLLIQLVLSFSDRSPLSETHDNPCESSASFLSRITFWITGLIVRGYRPLEGSD 240  
 241 LWSLNKEDTSQVVPVLVKNWKECAKTRKQPKVYVSSKDPAPKESKVDANEEVEAL 300  
 241 LWSLNKEDTSQVVPVLVKNWKECAKTRKQPKVYVSSKDPAPKESKVDANEEVEAL 300  
 301 IVKSPQKEWNPFLFKVLYKTEGYPFLMSFFFKAIHDLMMFSGPQLIKLIFVNDTKAPD 360  
 301 IVKSPQKEWNPFLFKVLYKTEGYPFLMSFFFKAIHDLMMFSGPQLIKLIFVNDTKAPD 360  
 361 WQGYEYTVLLFWTACLOTLVHOYFHCIFVSGMRKTAIVGAVRKALVITNSARKSSTV 420  
 361 WQGYEYTVLLFWTACLOTLVHOYFHCIFVSGMRKTAIVGAVRKALVITNSARKSSTV 420  
 421 GEIVNLSVDAQRFMDLATYINMIWSAPLOVILALYLWNLGFSVLGAVVWMLWVPVN 480  
 421 GEIVNLSVDAQRFMDLATYINMIWSAPLOVILALYLWNLGFSVLGAVVWMLWVPVN 480

QY 481 AVAMKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAWELAFKDKVLAIHQEELKVLK 540  
 DB 481 AVAMKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAWELAFKDKVLAIHQEELKVLK 540  
 QY 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVTTDENNILDQAQAFVLSALFNILRPPLNLP 600  
 DB 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVTTDENNILDQAQAFVLSALFNILRPPLNLP 600  
 QY 601 MVISSIVQASVSLAKRLRIFLSHEELEPDSIERRPVKDGGGTSITVRNATTTWARSDDPT 660  
 DB 601 MVISSIVQASVSLAKRLRIFLSHEELEPDSIERRPVKDGGGTSITVRNATTTWARSDDPT 660  
 QY 661 LINGITFSIPEGALVAVGVCGKSSLLSALLAEMDKVEGHVAIKGSVAYVPOQAWIQND 720  
 DB 661 LINGITFSIPEGALVAVGVCGKSSLLSALLAEMDKVEGHVAIKGSVAYVPOQAWIQND 720  
 QY 721 SLRENILFGCOLBEPYRSVIOACALLPDLBILPSGDRTEIGEGVNLSSGOKQKRVSLAR 780  
 DB 721 SLRENILFGCOLBEPYRSVIOACALLPDLBILPSGDRTEIGEGVNLSSGOKQKRVSLAR 780  
 QY 781 AVYSNADIYLPDDPLSAVDHVGKHFENVIGPKGMLKNKTRILVTHSMVSLPOVDVLIIV 840  
 DB 781 AVYSNADIYLPDDPLSAVDHVGKHFENVIGPKGMLKNKTRILVTHSMVSLPOVDVLIIV 840  
 QY 841 MSGGKISMGSYQELLARDGAFAEFLRTYASTEQDAEENGVTGVSQPGKEAKQMGNGM 900  
 DB 841 MSGGKISMGSYQELLARDGAFAEFLRTYASTEQDAEENGVTGVSQPGKEAKQMGNGM 900  
 QY 901 LVTSAGKQLQRLSSSSSYSGDISRHNNSTAEIQAFAKKEETWKLMEADKAQTQVKL 960  
 DB 901 LVTSAGKQLQRLSSSSSYSGDISRHNNSTAEIQAFAKKEETWKLMEADKAQTQVKL 960  
 QY 961 SVYWDYKAIGLFISFLFMCNHNVSALASNLWLSLWTDPIVNGTOEHTKVRLSVYG 1020  
 DB 961 SVYWDYKAIGLFISFLFMCNHNVSALASNLWLSLWTDPIVNGTOEHTKVRLSVYG 1020  
 QY 1021 ALGISQIAGVFGSMVAVSIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFSKEL 1080  
 DB 1021 ALGISQIAGVFGSMVAVSIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFSKEL 1080  
 QY 1081 DTVDSMIPEVIKPMGSLFNVIAGACIVILLATPIAAIIIPPLGLIYFFVQRYVASSRQL 1140  
 DB 1081 DTVDSMIPEVIKPMGSLFNVIAGACIVILLATPIAAIIIPPLGLIYFFVQRYVASSRQL 1140  
 QY 1141 KRLESVSRSPVYSHFNELTGLVSVIRAFEBQERFIHOSDLKVDENOKAYPSIVANRWLA 1200  
 DB 1141 KRLESVSRSPVYSHFNELTGLVSVIRAFEBQERFIHOSDLKVDENOKAYPSIVANRWLA 1200  
 QY 1201 VRLECVGNCIVLFAALFAVISRHSLSAGLVLSYSLSQVTTYLNWLVRMSSEMETNIVA 1260  
 DB 1201 VRLECVGNCIVLFAALFAVISRHSLSAGLVLSYSLSQVTTYLNWLVRMSSEMETNIVA 1260  
 QY 1261 VERLKVYSETEKEAPWQIOETAPPSSWPQVGRVFRNYCLRYREDLDFVLRIHINVTNGG 1320  
 DB 1261 VERLKVYSETEKEAPWQIOETAPPSSWPQVGRVFRNYCLRYREDLDFVLRIHINVTNGG 1320  
 QY 1321 EKVGIVGRTGAGKSLTGLFRINESAEGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
 DB 1321 EKVGIVGRTGAGKSLTGLFRINESAEGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
 QY 1381 SGSLRMNLDPPSQYSDSEVWTSLELAHDKDFVSALPKDLHDECAEGGENLSVGQRLVCL 1440  
 DB 1381 SGSLRMNLDPPSQYSDSEVWTSLELAHDKDFVSALPKDLHDECAEGGENLSVGQRLVCL 1440  
 QY 1441 ABALLRKTILVLDDEATAAVDLETDLLI QSTIRTOFEDCTVLTIAHRLNTIMDYTRIVL 1500  
 DB 1441 ABALLRKTILVLDDEATAAVDLETDLLI QSTIRTOFEDCTVLTIAHRLNTIMDYTRIVL 1500  
 QY 1501 DKGEIQBYGAPSDLLQORGLFYSMAXDAGLV 1531  
 DB 1501 DKGEIQBYGAPSDLLQORGLFYSMAXDAGLV 1531

RESULT 11  
 ADB92128  
 ID ADB92128 standard; protein; 1531 AA.  
 AC ADB92128;  
 XX  
 DT 04-DEC-2003 (first entry)  
 XX  
 DE Human MDR1 related protein sequence SEQ ID NO:678.  
 XX  
 KW irinotecan; colorectal cancer; cervical cancer; gastric cancer;  
 KW lung cancer; ovarian cancer; pancreatic cancer; malignant glioma;  
 KW multidrug resistance 1; MDR1; cytostatic; human; UGT1A1; MRP1; TOP1.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2003013535-A2.  
 XX  
 XX 20-FEB-2003.  
 XX  
 XX 23-JUL-2002; 2002WO-EP008220.  
 PF  
 XX 23-JUL-2001; 2001EP-00117608.  
 PR  
 XX 24-MAY-2002; 2002EP-00011710.  
 XX  
 PA (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.  
 XX  
 PI Heinrich G, Korb R;  
 XX  
 DR WPI; 2003-342400/32.  
 XX  
 PT New use of irinotecan for preparation of pharmaceutical compositions for  
 PT treating cancer in subject having genome with variant allele comprising  
 PT multidrug resistance 1 polynucleotide.  
 XX  
 PS Disclosure; SEQ ID NO 678; 104pp; English.  
 XX  
 CC The invention relates to a novel use of irinotecan or its derivative for  
 CC the preparation of a pharmaceutical composition for treating colorectal,  
 CC cervical, gastric, lung, ovarian or pancreatic cancer, or malignant  
 CC glioma in a subject having a genome with a variant allele which comprises  
 CC a multidrug resistance 1 (MDR1) polynucleotide. A composition of the  
 CC invention has cytostatic activity. The present sequence is used in the  
 CC exemplification of the invention.  
 XX  
 SQ Sequence 1531 AA;  
 Query Match 79.4%; Score 7860; DB 7; Length 1531;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MALRGFCADGSDPLNDNNTNTSNPDTKCFQNTVLVWVPCFYLWACFPYFLYLSRH 60  
 DB 1 MALRGFCADGSDPLNDNNTNTSNPDTKCFQNTVLVWVPCFYLWACFPYFLYLSRH 60  
 QY 61 DRGYIQTPLNKTALGFLMIVCWADLFYFWSRSGIFLAPVFLVSPILLGTTLLA 120  
 DB 61 DRGYIQTPLNKTALGFLMIVCWADLFYFWSRSGIFLAPVFLVSPILLGTTLLA 120  
 QY 121 TFLIQLERRKGVSQSGIMTLFWLVALCALAILRSKIMTALKEDAQVDLFRDITFVYVS 180  
 DB 121 TFLIQLERRKGVSQSGIMTLFWLVALCALAILRSKIMTALKEDAQVDLFRDITFVYVS 180  
 QY 181 LLLIQLVLSCFSDRSPFLSETHDNPCESSASFLSRITFWITGLIVRGVRLQEGSD 240  
 DB 181 LLLIQLVLSCFSDRSPFLSETHDNPCESSASFLSRITFWITGLIVRGVRLQEGSD 240  
 QY 241 LWSLNKEDTSEQVPLVKNWKECAKTRKQPVVYSSKDPAPQKSSKVDANEVEAL 300  
 DB 241 LWSLNKEDTSEQVPLVKNWKECAKTRKQPVVYSSKDPAPQKSSKVDANEVEAL 300  
 QY 301 IVKSPQKWNPSLFKVLTKTFGPFILMSFFFKAIHDLMMFSGPQILKLLIKFVNDTKAPD 360

DB 301 IVKSPQKWNPSLFKVLTKTFGPFILMSFFFKAIHDLMMFSGPQILKLLIKFVNDTKAPD 360  
 QY 361 WQGYFYTVLLFVTACLOTVLHQYFHICFVSGMRIKTAIVIGAVYRKALVITNSARKSTV 420  
 DB 361 WQGYFYTVLLFVTACLOTVLHQYFHICFVSGMRIKTAIVIGAVYRKALVITNSARKSTV 420  
 QY 421 GEIVNLSVDAQRPMDLATYINMIWSAPLOVILALYLLNLNLGSPSVLAGVAVMLMVPVN 480  
 DB 421 GEIVNLSVDAQRPMDLATYINMIWSAPLOVILALYLLNLNLGSPSVLAGVAVMLMVPVN 480  
 QY 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKLYAWELAFKDKVLAIRQEBELVKL 540  
 DB 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKLYAWELAFKDKVLAIRQEBELVKL 540  
 QY 541 KSAVLSAVGTFTWCTPFLVALCTFAVVVTIDENNILDAQTAFAFVSLALFNILRFPNLILP 600  
 DB 541 KSAVLSAVGTFTWCTPFLVALCTFAVVVTIDENNILDAQTAFAFVSLALFNILRFPNLILP 600  
 QY 601 MVISSIVOASVSLKRLRIFLSHEBLEPDSIERRPVKGGGTSITVRNATFTWARSDDPT 660  
 DB 601 MVISSIVOASVSLKRLRIFLSHEBLEPDSIERRPVKGGGTSITVRNATFTWARSDDPT 660  
 QY 661 LAGITFSIPEGALVAVVGVCGKSSLLSALLAEMDKVEGHVAKGVSAYVPPQAWIOND 720  
 DB 661 LAGITFSIPEGALVAVVGVCGKSSLLSALLAEMDKVEGHVAKGVSAYVPPQAWIOND 720  
 QY 721 SLRENILFGCOLBEPYRSVIOACALLPDLEILPSGDRTEIGEGKVNLSGGQKQVSLAR 780  
 DB 721 SLRENILFGCOLBEPYRSVIOACALLPDLEILPSGDRTEIGEGKVNLSGGQKQVSLAR 780  
 QY 781 AVYGNADIYLFDDPLSAVDHVGKHI FENVIGPKGMLKNKTRILVTHSMXYLPQVDVIIV 840  
 DB 781 AVYGNADIYLFDDPLSAVDHVGKHI FENVIGPKGMLKNKTRILVTHSMXYLPQVDVIIV 840  
 QY 841 MSGKISMSGSYQELLARDGAFABFLRTYASTEQDAEENGVTGVSQPGKEAKOMENGM 900  
 DB 841 MSGKISMSGSYQELLARDGAFABFLRTYASTEQDAEENGVTGVSQPGKEAKOMENGM 900  
 QY 901 LVTDAGKQLORQLSSSSSYSGDISRHNSHNSAEQKAEKKEETWKLMEADKAQGTQVKL 960  
 DB 901 LVTDAGKQLORQLSSSSSYSGDISRHNSHNSAEQKAEKKEETWKLMEADKAQGTQVKL 960  
 QY 961 SVYDYMKAIGLFTSFLSIFLFCMCHVSALASNYWLSLWTDPIVNGTQHTKVLRSVYG 1020  
 DB 961 SVYDYMKAIGLFTSFLSIFLFCMCHVSALASNYWLSLWTDPIVNGTQHTKVLRSVYG 1020  
 QY 1021 ALGISQGIATVFGYSMAVSI GIGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFSEL 1080  
 DB 1021 ALGISQGIATVFGYSMAVSI GIGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFSEL 1080  
 QY 1081 DTVDSMIPEVTKMFMSGLFNVI GACIVILLATPIAAIIIPGLGIYFFVQRFYVASSROL 1140  
 DB 1081 DTVDSMIPEVTKMFMSGLFNVI GACIVILLATPIAAIIIPGLGIYFFVQRFYVASSROL 1140  
 QY 1141 KRLESVSRSVPYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENOKAYPYSIVANRWLA 1200  
 DB 1141 KRLESVSRSVPYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENOKAYPYSIVANRWLA 1200  
 QY 1201 VRLCEVGNICIVLFAALFAVISRHLSAGLVLSVSYSLQVTTYLNMLVRMSEMETNIVA 1260  
 DB 1201 VRLCEVGNICIVLFAALFAVISRHLSAGLVLSVSYSLQVTTYLNMLVRMSEMETNIVA 1260  
 QY 1261 VERLKEYSETEKEAPWQIQETAPSSWPQVGRVEFRNYCLARYREDLDFVLRHINTVINGG 1320  
 DB 1261 VERLKEYSETEKEAPWQIQETAPSSWPQVGRVEFRNYCLARYREDLDFVLRHINTVINGG 1320  
 QY 1321 EKVGVGRTGAGKSLTLGLFRINESAGEEIIIDGINIAKIGLHDLRFKTIITIPDPVLF 1380  
 DB 1321 EKVGVGRTGAGKSLTLGLFRINESAGEEIIIDGINIAKIGLHDLRFKTIITIPDPVLF 1380  
 QY 1381 SGLSLMNLDPFSQVSDSEEVWTSLELAHLKDFVSALPDKLDHECAGGENLSVGQRLVCL 1440  
 DB 1381 SGLSLMNLDPFSQVSDSEEVWTSLELAHLKDFVSALPDKLDHECAGGENLSVGQRLVCL 1440

QY 1441 ARALLRKTILVDEATAVDLETDLIQSTIRTOPEDCVLTIAHRLNTIMDYTRVIL 1500  
 Db |||||||  
 QY 1441 ARALLRKTILVDEATAVDLETDLIQSTIRTOPEDCVLTIAHRLNTIMDYTRVIL 1500  
 Db |||||||  
 QY 1501 DKGEIQEYGAPSDLLQORGLFYSMARKDAGLV 1531  
 Db |||||||  
 QY 1501 DKGEIQEYGAPSDLLQORGLFYSMARKDAGLV 1531  
 Db |||||||

RESULT 12  
 ID ADD44764  
 XX ADD44764 standard; protein; 1531 AA.  
 AC ADD44764;  
 DT 29-JAN-2004 (first entry)  
 DE Human Protein P33527, SEQ ID NO 10193.  
 KW Human; pain; neuronal tissue; gene therapy;  
 KW spinal segmental nerve injury; chronic constriction injury; CCI;  
 KW spared nerve injury; SNI; Chung.  
 XX Homo sapiens.  
 XX WO2003016475-A2.  
 XX 27-FEB-2003.  
 XX 14-AUG-2002; 2002WO-US025765.  
 XX 14-AUG-2001; 2001US-0312147P.  
 PR 01-NOV-2001; 2001US-0345182P.  
 PR 26-NOV-2001; 2001US-0333347P.  
 XX (GEO ) GEN HOSPITAL CORP.  
 PA (FARB ) BAYER AG.  
 XX Woolf C, D'urso D, Befort K, Costigan M;  
 DR WPI; 2003-268312/26.  
 DR GENBANK; P33527.  
 PT New composition comprising two or more isolated polypeptides, useful for  
 PT preparing a medicament for treating pain in an animal.  
 XX Claim 1; Page; 1017pp; English.

The invention discloses a composition comprising two or more isolated rat or human polynucleotides or a polynucleotide which represents a fragment, derivative or allelic variation of the nucleic acid sequence. Also claimed are a vector comprising the novel polynucleotide, a host cell comprising the vector, a method for identifying a nucleotide sequence which is differentially regulated in an animal subjected to pain and a kit to perform the method, an array, a method for identifying an agent that increases or decreases the expression of the polynucleotide sequence that is differentially expressed in neuronal tissue of a first animal subjected to pain, a method for identifying a compound which regulates the expression of a polynucleotide sequence which is differentially expressed in an animal subjected to pain, a method for identifying a compound that regulates the activity of one or more of the polynucleotides, a method for producing a pharmaceutical composition, a method for identifying a compound or small molecule that regulates the activity in an animal of one or more of the polypeptides given in the specification, a method for identifying a compound useful in treating pain and a pharmaceutical composition comprising the one or more polypeptides or their antibodies. The polynucleotide or the compound that modulates its activity is useful for preparing a medicament for treating pain (e.g. spinal segmental nerve injury (Chung), chronic constriction injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene therapy). The sequence presented is a human protein (shown in Table 2 of the specification) which is differentially expressed during pain. Note:

CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 1531 AA;  
 Query Match 79.4%; Score 7860; DB 7; Length 1531;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MALRGFCSADGSDPLMDWNTWNTSNPDTKCFQNTVLVWVPCFYLWACFFPYFLYLSRH 60  
 Db |||||||  
 QY 1 MALRGFCSADGSDPLMDWNTWNTSNPDTKCFQNTVLVWVPCFYLWACFFPYFLYLSRH 60  
 Db |||||||  
 QY 61 DRGYIQTPLNKTALGFLWIVCWADLFFYSFWERSRGIFLAPVFLVSPLLGLITLLA 120  
 Db |||||||  
 QY 61 DRGYIQTPLNKTALGFLWIVCWADLFFYSFWERSRGIFLAPVFLVSPLLGLITLLA 120  
 Db |||||||  
 QY 121 TFLIQLERRKGVQSSGIMLTFWLVALCALAILRSKIMTALKEDAQVDLFRDITFYVYFS 180  
 Db |||||||  
 QY 121 TFLIQLERRKGVQSSGIMLTFWLVALCALAILRSKIMTALKEDAQVDLFRDITFYVYFS 180  
 Db |||||||  
 QY 181 LLLIQLVLSCFSDRSPLFSETHDNPCESSASFLSRITFWITGLIVRGYRQPLEGSD 240  
 Db |||||||  
 QY 181 LLLIQLVLSCFSDRSPLFSETHDNPCESSASFLSRITFWITGLIVRGYRQPLEGSD 240  
 Db |||||||  
 QY 241 LWSLNKEDTSEQVVPVLVKNWKKCAKTRQPKVYVSSKDPAPQKSSSKVDANEVEAL 300  
 Db |||||||  
 QY 241 LWSLNKEDTSEQVVPVLVKNWKKCAKTRQPKVYVSSKDPAPQKSSSKVDANEVEAL 300  
 Db |||||||  
 QY 301 IVKSPQKEWNPSPFKVLYKTFPGVFLMSFPFKAHDLMMFSGPOILKLIKFNVDTKAPD 360  
 Db |||||||  
 QY 301 IVKSPQKEWNPSPFKVLYKTFPGVFLMSFPFKAHDLMMFSGPOILKLIKFNVDTKAPD 360  
 Db |||||||  
 QY 361 WQGYFTVLLFVTAQLTLVHLQYFHCIFVSGMRKITAIVGAVYRKALVITNSARKSSTV 420  
 Db |||||||  
 QY 361 WQGYFTVLLFVTAQLTLVHLQYFHCIFVSGMRKITAIVGAVYRKALVITNSARKSSTV 420  
 Db |||||||  
 QY 421 GEIVNLSVDAQREMDLATYINMIWSAPLOVILALYLLNLNLSVLAGVAVWVLMVNVN 480  
 Db |||||||  
 QY 421 GEIVNLSVDAQREMDLATYINMIWSAPLOVILALYLLNLNLSVLAGVAVWVLMVNVN 480  
 Db |||||||  
 QY 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLYAMELAFKDKVLAIRQBELKVLK 540  
 Db |||||||  
 QY 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLYAMELAFKDKVLAIRQBELKVLK 540  
 Db |||||||  
 QY 541 KSAVLSAVGTFTWVCTPFLVALCTFAVYVITDENNILDQAOTAFVSLALFNILRPLNLP 600  
 Db |||||||  
 QY 541 KSAVLSAVGTFTWVCTPFLVALCTFAVYVITDENNILDQAOTAFVSLALFNILRPLNLP 600  
 Db |||||||  
 QY 601 MVTSISVQASVSLKRLRIFLSHEELEPDSIERRPVKGGGTNSITVRNATFTWARSPPPT 660  
 Db |||||||  
 QY 601 MVTSISVQASVSLKRLRIFLSHEELEPDSIERRPVKGGGTNSITVRNATFTWARSPPPT 660  
 Db |||||||  
 QY 661 LMGITFIPGALVAVVGVQVCGKSSLLSALLAEMDKVEGHVAIKGSVAVVPOQAWTQND 720  
 Db |||||||  
 QY 661 LMGITFIPGALVAVVGVQVCGKSSLLSALLAEMDKVEGHVAIKGSVAVVPOQAWTQND 720  
 Db |||||||  
 QY 721 SLRNLIFGQLEBEPYRVSIVIQACALIPDLIELIPSGDRTEIGEGVNLSSGQKQVSLAR 780  
 Db |||||||  
 QY 721 SLRNLIFGQLEBEPYRVSIVIQACALIPDLIELIPSGDRTEIGEGVNLSSGQKQVSLAR 780  
 Db |||||||  
 QY 781 AVYSNADILYFDPLSAVDHVGKHFENVIQKGMKLNKTRILVTHSMVLPQVDVILV 840  
 Db |||||||  
 QY 781 AVYSNADILYFDPLSAVDHVGKHFENVIQKGMKLNKTRILVTHSMVLPQVDVILV 840  
 Db |||||||  
 QY 841 MSGGKISEMSYQELLARDGAFELFTYASTEQEQDAEENGVTGVSQPKGKAKOMENGM 900  
 Db |||||||  
 QY 841 MSGGKISEMSYQELLARDGAFELFTYASTEQEQDAEENGVTGVSQPKGKAKOMENGM 900  
 Db |||||||  
 QY 901 LVTDGAKQLQRLSSSSSSSGDISRHNSHTAELOKAEAKKETWKLMEADKAQTQGVKL 960  
 Db |||||||  
 QY 901 LVTDGAKQLQRLSSSSSSSGDISRHNSHTAELOKAEAKKETWKLMEADKAQTQGVKL 960  
 Db |||||||

QY 961 SVYDYMKAIGLFTISFLSIFLPMCHVNSALASNYWLSLWTDPIVNGTOEHTKVLRSYVG 1020  
 DB 961 SVYDYMKAIGLFTISFLSIFLPMCHVNSALASNYWLSLWTDPIVNGTOEHTKVLRSYVG 1020  
 QY 1021 ALGISQIGTAVGYSMAVSGIGGLASRCLHVDLLHSILRSPMSFFERTSGNLVNRFSKEL 1080  
 DB 1021 ALGISQIGTAVGYSMAVSGIGGLASRCLHVDLLHSILRSPMSFFERTSGNLVNRFSKEL 1080  
 QY 1081 DTVDSMIEVIMKMGSLFNIVGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQL 1140  
 DB 1081 DTVDSMIEVIMKMGSLFNIVGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQL 1140  
 QY 1141 KRLESVSRSPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENQKAYPSIVANRWLA 1200  
 DB 1141 KRLESVSRSPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENQKAYPSIVANRWLA 1200  
 QY 1201 VRLECVGNCIVLFAALFAVISHRSLSAGLVGLSVSYSLQVTTYLNWLVMSSEMETNIVA 1260  
 DB 1201 VRLECVGNCIVLFAALFAVISHRSLSAGLVGLSVSYSLQVTTYLNWLVMSSEMETNIVA 1260  
 QY 1261 VERLKEYSETEKAPQIETAPPSSWQVGRVEFRNCLRYREDLDFVLRHINVTINGG 1320  
 DB 1261 VERLKEYSETEKAPQIETAPPSSWQVGRVEFRNCLRYREDLDFVLRHINVTINGG 1320  
 QY 1321 EKVGIVRTGACKSSLTGLFRINESAEGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
 DB 1321 EKVGIVRTGACKSSLTGLFRINESAEGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
 QY 1381 SGLSRMNLDPFSQYSDVEVWTSLELAHLKDFVSALPKLDHECARGGENLSVGQRQLVCL 1440  
 DB 1381 SGLSRMNLDPFSQYSDVEVWTSLELAHLKDFVSALPKLDHECARGGENLSVGQRQLVCL 1440  
 QY 1441 ARALARKTILVLDATAVDETDLLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 DB 1441 ARALARKTILVLDATAVDETDLLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 QY 1501 DKGEIQEYCAPSLLQORGLFYSMADAGLV 1531  
 DB 1501 DKGEIQEYCAPSLLQORGLFYSMADAGLV 1531  
 RESULT 13  
 ADN97111  
 ID ADN97111 standard; protein; 1531 AA.  
 AC ADN97111;  
 XX  
 DT 01-JUL-2004 (first entry)  
 DE  
 DE MRP1 protein.  
 XX  
 KW multiple drug resistance protein; MRP; Drosophila melanogaster;  
 KW Anopheles gambiae; insecticide.  
 XX  
 OS Homo sapiens.  
 XX  
 FN WO2004029088-A2.  
 XX  
 PD 08-APR-2004.  
 XX  
 PF 25-SEP-2003; 2003WO-EP012400.  
 XX  
 PR 26-SEP-2002; 2002US-0413469P.  
 XX  
 XX (INSP ) INST PASTEUR.  
 PA (CNRS ) CENT NAT RECH SCI.  
 XX  
 PI Roth CW, Brey PT, Holm I, Graillies M, Rzhetsky A;  
 XX  
 DR WPI; 2004-305150/28.  
 XX  
 PT New polynucleotide sequence encoding multiple drug resistance proteins

PT from Drosophila melanogaster or Anopheles gambiae, useful in developing  
 PT effective insecticides.  
 XX Claim 1; SEQ ID NO 6; 58pp; English.  
 XX  
 CC The present invention relates to a purified polynucleotide or its  
 CC fragment and comprises a sequence encoding multiple drug resistance  
 CC proteins (MRPs) from Drosophila melanogaster or Anopheles gambiae. The  
 CC polynucleotide is useful in developing effective insecticides. The  
 CC present sequence represents human MRP1.  
 XX  
 SQ Sequence 1531 AA;

Query Match 79.4%; Score 7860; DB 8; Length 1531;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCSADGSDPLMDWNVTWNTSNPDFTKCFQNTVLVWVPCFVLMACFPFYLYLSRH 60  
 DB 1 MALRGFCSADGSDPLMDWNVTWNTSNPDFTKCFQNTVLVWVPCFVLMACFPFYLYLSRH 60  
 QY 61 DRGVIQMTPLNKTKTALGFLLWVWADLFYSFWSRSGIFLAPVFLVSPILLGTTLLA 120  
 DB 61 DRGVIQMTPLNKTKTALGFLLWVWADLFYSFWSRSGIFLAPVFLVSPILLGTTLLA 120  
 QY 121 TFLIQLERRKGVQSSGIMLTFWLVALVCAALAILRSKIMTALKEDAQVDFRDITFYVYFS 180  
 DB 121 TFLIQLERRKGVQSSGIMLTFWLVALVCAALAILRSKIMTALKEDAQVDFRDITFYVYFS 180  
 QY 181 LLLIQLVLSLCSFSDRSPLSETHIDNPSPSSASFLSRITFWITGLIVRGVRLQEGSD 240  
 DB 181 LLLIQLVLSLCSFSDRSPLSETHIDNPSPSSASFLSRITFWITGLIVRGVRLQEGSD 240  
 QY 241 LWSLNKEDTSQVVPVLVKNWKKCAKTRKOPKVVYSSKDPAPQPKESKVDANEVEAL 300  
 DB 241 LWSLNKEDTSQVVPVLVKNWKKCAKTRKOPKVVYSSKDPAPQPKESKVDANEVEAL 300  
 QY 301 IVKSPQKEWNPFLFKVLYKTFPGYPLMSFFFKAIHDLMMFSGPQILKLLIKFVNDTKAPD 360  
 DB 301 IVKSPQKEWNPFLFKVLYKTFPGYPLMSFFFKAIHDLMMFSGPQILKLLIKFVNDTKAPD 360  
 QY 361 WQGYFYTVLLFVTTACLTQLVHLHOYFHICFVSGMRITKAVIGAVYRKALVITNSARKSSTV 420  
 DB 361 WQGYFYTVLLFVTTACLTQLVHLHOYFHICFVSGMRITKAVIGAVYRKALVITNSARKSSTV 420  
 QY 421 GEIVNLSVDAQRFMDLATYINMIWSAPLOVTLALYLWNLGSPVLGAVVWVLMVFN 480  
 DB 421 GEIVNLSVDAQRFMDLATYINMIWSAPLOVTLALYLWNLGSPVLGAVVWVLMVFN 480  
 QY 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAWELAFKDKVLAIROEELKVLK 540  
 DB 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAWELAFKDKVLAIROEELKVLK 540  
 QY 541 KSAVLSAVGTTWCTPPLVALCTPAVVTVDENNILDAQTAFAVSLALFNILRFPNLILP 600  
 DB 541 KSAVLSAVGTTWCTPPLVALCTPAVVTVDENNILDAQTAFAVSLALFNILRFPNLILP 600  
 QY 601 MVISSIVQASVSLKELRIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWASDPT 660  
 DB 601 MVISSIVQASVSLKELRIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWASDPT 660  
 QY 661 LINGITFISPEGALVAVGVQCGKSSLLSALLAEMDKVEGHVAIKGSVAVYVQQAWIQND 720  
 DB 661 LINGITFISPEGALVAVGVQCGKSSLLSALLAEMDKVEGHVAIKGSVAVYVQQAWIQND 720  
 QY 721 SLRENILFGCOLEPEYRSVIOACALLPDLEILPSGDRTEIGEKGVNLGGQKQVSLAR 780  
 DB 721 SLRENILFGCOLEPEYRSVIOACALLPDLEILPSGDRTEIGEKGVNLGGQKQVSLAR 780  
 QY 781 AVYSNADIYLFDDPLSAVDHVGKHI FENVIGPKMLXNKTRILVTHSNLYLPQVDVILV 840  
 DB 781 AVYSNADIYLFDDPLSAVDHVGKHI FENVIGPKMLXNKTRILVTHSNLYLPQVDVILV 840

```
QY 841 MSGKISMGSGYQELLARDGAFABELRTYASTEQDQDAEENGVTGSGPGKEAKOMENG 900
Db 841 MSGKISMGSGYQELLARDGAFABELRTYASTEQDQDAEENGVTGSGPGKEAKOMENG 900
QY 901 LVTSAGKQORQLSSSSSYSGDISRHNSHNSAELOKABAKYEEWKLMEADKAQTGVKL 960
Db 901 LVTSAGKQORQLSSSSSYSGDISRHNSHNSAELOKABAKYEEWKLMEADKAQTGVKL 960
QY 961 SVYDYMKAIGLFISFLSIFLPMCNHVSALASNYWLSLWTDPIVNGTQEHKVRLSVYG 1020
Db 961 SVYDYMKAIGLFISFLSIFLPMCNHVSALASNYWLSLWTDPIVNGTQEHKVRLSVYG 1020
QY 1021 ALGISQIAGVGSMAVIGIGLASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFSEL 1080
Db 1021 ALGISQIAGVGSMAVIGIGLASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFSEL 1080
QY 1081 DTVDMSIPEVIMFMGSLFNIVIGACIVILLATPIAAIIIPPLGLIYFFVQRYVASSRQL 1140
Db 1081 DTVDMSIPEVIMFMGSLFNIVIGACIVILLATPIAAIIIPPLGLIYFFVQRYVASSRQL 1140
QY 1141 KRLESVSPVYSHENETLLGVSVIRAFEPQERFTHQSDLVKVDENQKAYPSIVANRWLA 1200
Db 1141 KRLESVSPVYSHENETLLGVSVIRAFEPQERFTHQSDLVKVDENQKAYPSIVANRWLA 1200
QY 1201 VRLCVCNCIVLFAALFAVIRSRHLSAGLVLSYSLSQVTTYLNWLVRMSEMETNIVA 1260
Db 1201 VRLCVCNCIVLFAALFAVIRSRHLSAGLVLSYSLSQVTTYLNWLVRMSEMETNIVA 1260
QY 1261 VERLUKEYETKEAPWQIQTAPPSSWPQGRVFRNVCYLRVEDLDPVLRHINVTINGG 1320
Db 1261 VERLUKEYETKEAPWQIQTAPPSSWPQGRVFRNVCYLRVEDLDPVLRHINVTINGG 1320
QY 1321 EKVGIVGRTGAGKSLTGLFRINESARGEIIIDGINIAKIGLHDLRPKITIIIPQDPVLF 1380
Db 1321 EKVGIVGRTGAGKSLTGLFRINESARGEIIIDGINIAKIGLHDLRPKITIIIPQDPVLF 1380
QY 1381 SGSLRMNLDPFSQYSDSEVWTSLELAHLKDFVSALPDKLDHECAEGGNLSVGQRQLVCL 1440
Db 1381 SGSLRMNLDPFSQYSDSEVWTSLELAHLKDFVSALPDKLDHECAEGGNLSVGQRQLVCL 1440
QY 1441 ARALLRKTKILVDEATAVLETDLLIQSTIRTOQFEDCTVLTIAHRLNTIMDYTRVIL 1500
Db 1441 ARALLRKTKILVDEATAVLETDLLIQSTIRTOQFEDCTVLTIAHRLNTIMDYTRVIL 1500
QY 1501 DKGEIQEYGAQSDLLQQRGLFYSMAKDAGLV 1531
Db 1501 DKGEIQEYGAQSDLLQQRGLFYSMAKDAGLV 1531

RESULT 14
ADN95929
ID ADN95929 standard; protein; 1530 AA.
XX AC ADN95929;
XX DT
XX DE Human BEC/LEC-related protein sequence SeqID853.
XX
KW growth; differentiation; blood endothelial cell; BEC;
KW lymphatic endothelial cell; LEC; hereditary lymphoedema; VEGFR-3;
KW lymphatic growth agent; VEGF-C; VEGF-D; antiangiogenic; cytostatic;
KW vasotrophic; antiinflammatory; gene therapy; endothelial cell disorder;
KW inflammatory disease; cancer metastasis; lymphatic system; human.
XX
OS Homo sapiens.
XX
FN WO2003080640-A1.
XX
PD 02-OCT-2003.
XX
PF 07-MAR-2003; 2003WO-US006900.
XX
```

```
PR 07-MAR-2002; 2002US-0363019P.
XX (LUDW-) LUDWIG INST CANCER RES.
PA (LICN) LICENTIA LTD.
XX
XX Alitalo K, Makinen T, Petrova T, Saharinen P, Saharinen J;
XX WPI: 2003-876899/81.
XX N-PSDB; ADN95930.
XX
XX Example 1; SEQ ID NO 853; 176pp; English.
```

This invention relates to a method of differentially modulating the growth or differentiation of blood endothelial cells (BEC) or lymphatic endothelial cells (LEC) comprises contacting endothelial cells with a composition comprising an agent that differentially modulates blood or lymphatic endothelial cells. Treating hereditary lymphoedema comprises identifying a human subject with lymphoedema and with a mutation in at least one allele of a gene encoding a LEC protein, where the mutation correlates with lymphoedema in human subjects, and with the proviso that the LEC protein is not VEGFR-3; and administering to the subject a composition comprising a lymphatic growth agent selected from VEGF-C or VEGF-D polypeptides and polynucleotides. The invention may be useful for the development of compounds with an antiangiogenic, cytostatic, vasotrophic or antiinflammatory activity or for gene therapy. The method is useful in modulating the growth or differentiation of blood endothelial cells or lymphatic endothelial cells, in treating hereditary lymphoedema, in screening for an endothelial cell disorder or predisposition to the disorder or in monitoring the efficacy or toxicity of a drug on endothelial cells. The agent is useful in manufacturing a medicament for the differential modulation of blood vessel endothelial cell or lymphatic vessel endothelial cell growth or differentiation. The lymphatic growth agent may also be used in manufacturing a medicament for the treatment of hereditary lymphoedema resulting from a mutation in a LEC gene or of other diseases involving the lymphatic vessels, such as various inflammatory diseases and cancer metastasis via the lymphatic system. The present sequence is that of a human LEC/BEC differentially expressed protein which is related to the method of the invention. Note: This sequence does not appear in the specification but was obtained by the indexer using the source data given in table 14 of the specification.

Seq Sequence 1530 AA;

Query Match 79.3%; Score 7856; DB 7; Length 1530;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1530; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MALRGFCSADGSDPLWDNVTWNTSNPDFTKCFONTVLVWVPCFYLMACFPFFLYLSRH 60
Db 1 MALRGFCSADGSDPLWDNVTWNTSNPDFTKCFONTVLVWVPCFYLMACFPFFLYLSRH 60
QY 61 DRGYIOMTPLNKTALGFLLIWVCWADLVFSFWSRGIPLAPVFLVSPILLGITLLA 120
Db 61 DRGYIOMTPLNKTALGFLLIWVCWADLVFSFWSRGIPLAPVFLVSPILLGITLLA 120
QY 121 TFLIQLERRKGVSQSSGIMLTFWLVALCALAILRSKIMTALKEDAQVDLPFDITFYVYFS 180
Db 121 TFLIQLERRKGVSQSSGIMLTFWLVALCALAILRSKIMTALKEDAQVDLPFDITFYVYFS 180
QY 181 LLLIQLVLSCFSDRSPLFSETHIDPNPCPESSASFLSRITFWWTITGLIVRGYRQPLEGSD 240
Db 181 LLLIQLVLSCFSDRSPLFSETHIDPNPCPESSASFLSRITFWWTITGLIVRGYRQPLEGSD 240
QY 241 LWSLNKEDTSEQVVPVLVKNWKKCAKTRQPKVYVSSKDPQPKSSKVDANEEVEAL 300
Db 241 LWSLNKEDTSEQVVPVLVKNWKKCAKTRQPKVYVSSKDPQPKSSKVDANEEVEAL 300
QY 301 IVKSPQKEWNPSLFKVLYKTFGPFYFLMSFFFKALHDLMMFSGPQILKLLIKFVNDTKAPD 360
Db 301 IVKSPQKEWNPSLFKVLYKTFGPFYFLMSFFFKALHDLMMFSGPQILKLLIKFVNDTKAPD 360
QY 361 WQGYFTYVLLFVTACLOTVLVHQVFIHCFVSGMRKTAIVTGANVRKALVITNSARKSSTV 420
Db 361 WQGYFTYVLLFVTACLOTVLVHQVFIHCFVSGMRKTAIVTGANVRKALVITNSARKSSTV 420
```

Db 361 WGGYFYTVLLFTVACQLVLHQQYFHCIFVSGMRITKTAIVGAVYRKALVITNSARKSSTV 420  
Qy 421 GEIVNLSYDAQFMDLATYINMIWSAPLQVILALYLWLNLPSPVLAGVAVMLVMPVN 480  
Db 421 GEIVNLSYDAQFMDLATYINMIWSAPLQVILALYLWLNLPSPVLAGVAVMLVMPVN 480  
Qy 481 AVAMKTKTYQVAHKKSDNRKLMNEILNGIKVLKYAWELAFKDKVLAIROEBELKVLK 540  
Db 481 AVAMKTKTYQVAHKKSDNRKLMNEILNGIKVLKYAWELAFKDKVLAIROEBELKVLK 540  
Qy 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTDENNILDAQTAFAVSLALFNILRFPNLILP 600  
Db 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTDENNILDAQTAFAVSLALFNILRFPNLILP 600  
Qy 601 MVISSIVQASVSLKRLIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWASDPPT 660  
Db 601 MVISSIVQASVSLKRLIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWASDPPT 660  
Qy 661 LINGITFSIPEGALVAVGVGCGKSSLLSALLAEMDKVEGHVAIKGSVAVYVQQAWIQND 720  
Db 661 LINGITFSIPEGALVAVGVGCGKSSLLSALLAEMDKVEGHVAIKGSVAVYVQQAWIQND 720  
Qy 721 SIRENILFGCOLLEPYRSVIOACALLPDLLEILPSGDRTEIGKGVNLSGGOKQKVSRLAR 780  
Db 721 SIRENILFGCOLLEPYRSVIOACALLPDLLEILPSGDRTEIGKGVNLSGGOKQKVSRLAR 780  
Qy 781 AVYSNADIYLFDDPLSADVAHVGHKIFENVIGPKGMLKNKTRILVTHSMSYLPQVDVILV 840  
Db 781 AVYSNADIYLFDDPLSADVAHVGHKIFENVIGPKGMLKNKTRILVTHSMSYLPQVDVILV 840  
Qy 841 MSGGKISEMGSYQELLARDGAPAEFLRTYASTEQDAEENGVTGSGFGKEAKOMENG 900  
Db 841 MSGGKISEMGSYQELLARDGAPAEFLRTYASTEQDAEENGVTGSGFGKEAKOMENG 900  
Qy 901 LVTDGAGLQRLQSSSSSYSGDISRHNSTAELOKAEAKKEETWKLMEADKAQGTQVKL 960  
Db 901 LVTDGAGLQRLQSSSSSYSGDISRHNSTAELOKAEAKKEETWKLMEADKAQGTQVKL 960  
Qy 961 SVYDYMKAIGLFIISFLGIFPMCNHVSALASNYWLSLWTDPIVNGTQHTKVRLSVYG 1020  
Db 961 SVYDYMKAIGLFIISFLGIFPMCNHVSALASNYWLSLWTDPIVNGTQHTKVRLSVYG 1020  
Qy 1021 ALGISQGIAGVGYSMASVIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVRFSKEL 1080  
Db 1021 ALGISQGIAGVGYSMASVIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVRFSKEL 1080  
Qy 1081 DTVDMSIPEVIMKFMGSLFNIVGACIVILLATPIAAIIIPDLGLIYFPQFPYVASSRQL 1140  
Db 1081 DTVDMSIPEVIMKFMGSLFNIVGACIVILLATPIAAIIIPDLGLIYFPQFPYVASSRQL 1140  
Qy 1141 KRLESVSRSPVYSHFNETLLGVSVIRAFEEQERFTHQSDLKVDENOKAYYPSIVANRWLA 1200  
Db 1141 KRLESVSRSPVYSHFNETLLGVSVIRAFEEQERFTHQSDLKVDENOKAYYPSIVANRWLA 1200  
Qy 1201 VRLECVGNCIVLFAALFAVIRSHSLSAGLVGLSVSYSLQVTTYLNLWRMSSEMETNIVA 1260  
Db 1201 VRLECVGNCIVLFAALFAVIRSHSLSAGLVGLSVSYSLQVTTYLNLWRMSSEMETNIVA 1260  
Qy 1261 VERLKEYSETEKAPWQIQTAPPSPWPQVGRVEFRNYCLRYREDLDFVLRHINVTINGG 1320  
Db 1261 VERLKEYSETEKAPWQIQTAPPSPWPQVGRVEFRNYCLRYREDLDFVLRHINVTINGG 1320  
Qy 1321 EKVGIVGRTGAGKSSLTGLFRINSEAGEIIIDGINIAKILGHDLRPFKTIIPDPVLF 1380  
Db 1321 EKVGIVGRTGAGKSSLTGLFRINSEAGEIIIDGINIAKILGHDLRPFKTIIPDPVLF 1380  
Qy 1381 SGLSRMNLDPFSQYDEEVEWTSLELAHLKDFVSALPDKLDHECAGGENLSVGQRQLVCL 1440  
Db 1381 SGLSRMNLDPFSQYDEEVEWTSLELAHLKDFVSALPDKLDHECAGGENLSVGQRQLVCL 1440  
Qy 1441 ARALLRKTKILVLDATAVLDLTDLLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
Db 1441 ARALLRKTKILVLDATAVLDLTDLLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500

Qy 1501 DKGEIQEYGAPSDLLQQRGLFYSMKADAGL 1530  
Db 1501 DKGEIQEYGAPSDLLQQRGLFYSMKADAGL 1530

## RESULT 15

AARS4928  
ID AARS4928 standard; protein; 1531 AA.

XX AARS4928;

AC AARS4928;

XX 25-MAR-2003 (revised)

DT 14-OCT-1994 (first entry)

XX Multidrug resistance protein.

XX Multidrug resistance protein; MRP; H69AR; cancer cell line; stem cell;

KW cardiac muscle; transgenic animal.

XX Homo sapiens.

OS

PN WO9410303-A1.

XX PD 11-MAY-1994.

XX PF 27-OCT-1993; 93WO-CA000439.

XX PR 27-OCT-1992; 92US-00966923.

XX PR 08-MAR-1993; 93US-00029340.

XX PA (TOOH) UNIV QUEBENS KINGSTON.

XX PI Deesley RG, Cole SPC;

XX DR WPI; 1994-167460/20.

XX DR N-PSDB; AAQ65377.

XX PT Multi-drug resistance gene - encodes protein capable of conferring multi-

PT drug resistance on cells, useful in diagnostic and treatment methods.

XX PS Disclosure; Page 69-74; 101pp; English.

XX CC The multidrug resistant cancer cell line H69AR (ATCC CRL 11350) was used

CC to identify cDNA encoding a novel protein associated with multidrug

CC resistance, MRP. MRP may be expressed in e.g. hematopoietic stem cells or

CC cardiac muscle, or in transgenic animals, or can be used to raise

CC antibodies. (Updated on 25-MAR-2003 to correct PN field.)

XX SQ Sequence 1531 AA;

Query Match 79.3%; Score 7849; DB 2; Length 1531;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MALRGFCSADGSDPLDNDWNTWNTSNPDFTKCFONTVLVWVPCFYLWACFPFYLYLSRH 60  
Db 1 MALRGFCSADGSDPLDNDWNTWNTSNPDFTKCFONTVLVWVPCFYLWACFPFYLYLSRH 60

Qy 61 DRGYIQMTPLNKTALGFLWLWVWADLFPYSFWERSRGIFLAPVFLVSPDLLGTTLLA 120  
Db 61 DRGYIQMTPLNKTALGFLWLWVWADLFPYSFWERSRGIFLAPVFLVSPDLLGTTLLA 120

Qy 121 TFLQLERRKGVQSSGIMLTFWLVALCALAILSKINTALKEDAQVDLFRDITFYVYFS 180  
Db 121 TFLQLERRKGVQSSGIMLTFWLVALCALAILSKINTALKEDAQVDLFRDITFYVYFS 180

Qy 181 LLLQLVLSCFSDRSPFLFSETIHDPNPCPSSASFLSRITFWITGLIVRGVROPLEGSD 240  
Db 181 LLLQLVLSCFSDRSPFLFSETIHDPNPCPSSASFLSRITFWITGLIVRGVROPLEGSD 240

Qy 241 LWSLNKEDTSSQVVPVLVKNWKECAKTRKQPVKVYSSKDPAPQKSSKVDANSEVAL 300  
Db 241 LWSLNKEDTSSQVVPVLVKNWKECAKTRKQPVKVYSSKDPAPQKSSKVDANSEVAL 300

Db 241 LWSLNKEDTSEQVPLVKNWKKCECAKTRKQPVKVYSSKDPAPKESKSDVADANEVEAL 300  
 Qy 301 IVKSPKEMNSLKFVLYKTPGPFELMSFFPKAHLDLAMFSGPOLIKLLIKFVNDTKAPD 360  
 Db 301 IVKSPKEMNSLKFVLYKTPGPFELMSFFPKAHLDLAMFSGPOLIKLLIKFVNDTKAPD 360  
 Qy 361 WQGYFYTVLLFTACLOTVLHQYFPHICFVSGMRKTAIVIGAVYRKALVITNSARKSSTV 420  
 Db 361 WQGYFYTVLLFTACLOTVLHQYFPHICFVSGMRKTAIVIGAVYRKALVITNSARKSSTV 420  
 Qy 421 GEIVNLSVDAQRFMDLTYINIMINSAPLOVILLYLNLNLGSPSVLAGVAVMLVMPVN 480  
 Db 421 GEIVNLSVDAQRFMDLTYINIMINSAPLOVILLYLNLNLGSPSVLAGVAVMLVMPVN 480  
 Qy 481 AVMAWKTQYQVAHMKSKDNRIKLMNEILNGIKVLYAWELAFKDKVLAIROELKVLK 540  
 Db 481 AVMAWKTQYQVAHMKSKDNRIKLMNEILNGIKVLYAWELAFKDKVLAIROELKVLK 540  
 Qy 541 KSAYLSAVGTFTWCTPFLVALCTPAVYVTTIDENNILDAQTAFVSLALFNILRPPLNLP 600  
 Db 541 KSAYLSAVGTFTWCTPFLVALCTPAVYVTTIDENNILDAQTAFVSLALFNILRPPLNLP 600  
 Qy 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGGTNSITVRNATFTWARSDDPT 660  
 Db 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGGTNSITVRNATFTWARSDDPT 660  
 Qy 661 LMGITFTSPEGALVAVVGQCGKSLLSALLAEMDKVEGHVAKGVSAYVPOQAWIOND 720  
 Db 661 LMGITFTSPEGALVAVVGQCGKSLLSALLAEMDKVEGHVAKGVSAYVPOQAWIOND 720  
 Qy 721 SLRENILFGCOLLEPYYSVIOACALLPDLLEIPSGDRTEIGEGKVNLSGGQKQVSLAR 780  
 Db 721 SLRENILFGCOLLEPYYSVIOACALLPDLLEIPSGDRTEIGEGKVNLSGGQKQVSLAR 780  
 Qy 781 AVYSNADIYLFDDPLSADVAHVGHKHFENVIGPKGMLKNKTRILVTHSMSYLPQVDVIV 840  
 Db 781 AVYSNADIYLFDDPLSADVAHVGHKHFENVIGPKGMLKNKTRILVTHSMSYLPQVDVIV 840  
 Qy 841 MSGGKISEMGSYQELLARDGAPAEFLRYASTEQEODAEENGVTGVSFGPKGKAKOMENG 900  
 Db 841 MSGGKISEMGSYQELLARDGAPAEFLRYASTEQEODAEENGVTGVSFGPKGKAKOMENG 900  
 Qy 901 LVTDSAGKQLQRLSSSSSYSGDISRHHNSTAELQAEAKKEETWKLMEADKAQTQGVKL 960  
 Db 901 LVTDSAGKQLQRLSSSSSYSGDISRHHNSTAELQAEAKKEETWKLMEADKAQTQGVKL 960  
 Qy 961 SVYWDYMKAIGLFISFLIFPMCHNVHVSALASNYWLSLWTDPIVNGTQEHKVRLSVYG 1020  
 Db 961 SVYWDYMKAIGLFISFLIFPMCHNVHVSALASNYWLSLWTDPIVNGTQEHKVRLSVYG 1020  
 Qy 1021 ALGISQGIATVFGYSMAVSIIGGILASRCLHVLDLHLSILRSPMSFFERTPSGNLVNRFKSEL 1080  
 Db 1021 ALGISQGIATVFGYSMAVSIIGGILASRCLHVLDLHLSILRSPMSFFERTPSGNLVNRFKSEL 1080  
 Qy 1081 DTVDSMIPEVIKMFMSGLFNIVGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQL 1140  
 Db 1081 DTVDSMIPEVIKMFMSGLFNIVGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQL 1140  
 Qy 1141 KLESVSRSPPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENQKAYPSIVANRWLA 1200  
 Db 1141 KLESVSRSPPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENQKAYPSIVANRWLA 1200  
 Qy 1201 VRLECVGNCIVLFAALFAVISHRSLSAGLVLSVSYSLQVTTYLNWLVMSSEMETNIVA 1260  
 Db 1201 VRLECVGNCIVLFAALFAVISHRSLSAGLVLSVSYSLQVTTYLNWLVMSSEMETNIVA 1260  
 Qy 1261 VERLKEYSETEKAPWQIQETAPPSPWPQGVREFRNYCLARYREDLDFVLRHINVTINGG 1320  
 Db 1261 VERLKEYSETEKAPWQIQETAPPSPWPQGVREFRNYCLARYREDLDFVLRHINVTINGG 1320  
 Qy 1321 EXKGVIGRTGAGKSSLTGLFRINESARGEIIIDGINIAKIGLHDLRFXKTIIPQDPVLF 1380  
 Db 1321 EXKGVIGRTGAGKSSLTGLFRINESARGEIIIDGINIAKIGLHDLRFXKTIIPQDPVLF 1380

Qy 1381 SGSLRMNLDPFQSYSDSEVWTSLELAHLKDFVSALPKLDHECAEGGENLSVGORQLVCL 1440  
 Db 1381 SGSLRMNLDPFQSYSDSEVWTSLELAHLKDFVSALPKLDHECAEGGENLSVGORQLVCL 1440  
 Qy 1441 ARALLRKTILVLDEATAAVDLETTDDLIQSTIRTFQFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 Db 1441 ARALLRKTILVLDEATAAVDLETTDDLIQSTIRTFQFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 Qy 1501 DKGEIOEYGAPSDILQOQGLFYSMAXDAGLV 1531  
 Db 1501 DKGEIOEYGAPSDILQOQGLFYSMAXDAGLV 1531

Search completed: March 18, 2005, 11:02:13  
 Job time : 147.282 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2005, 10:57:20 ; Search time 45.2651 Seconds  
(without alignments)  
4096.085 Million cell updates/sec

Title: US-10-665-283-1  
Perfect score: 9903  
Sequence: 1 MALRGFCSADGSDPLMDWNV.....RSVAVAKPKFSIPDLSL 1927

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_79:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7849	79.3	1531	1 DVHUAR	multidrug resistan
2	4484.5	45.3	1527	2 JE0336	canalicular multis
3	3587	36.2	1545	1 S71841	multidrug resistan
4	3507.5	35.4	1541	1 S71839	canalicular multid
5	3289	33.2	1494	2 E89447	protein F57C12.4 [
6	3225	32.6	1573	2 T21219	hypothetical prote
7	3195.5	32.3	1502	2 T42216	multidrug resistan
8	2838	28.7	1515	1 S51863	cadmium resistance
9	2732	27.6	1398	2 T20434	hypothetical prote
10	2548	25.7	1478	2 T38712	ABC transporter SP
11	2384.5	24.1	1623	2 T01369	ABC transporter At
12	2356.5	23.8	1622	2 D86428	glutathione S-conj
13	2326	23.5	1495	2 E86428	probable ABC trans
14	2312	23.3	1144	2 T27408	hypothetical prote
15	2294.5	23.2	1559	1 S64757	probable membrane
16	2263.5	22.9	1488	2 F86428	probable ABC trans
17	2238.5	22.6	1516	2 F84919	glutathione-conjug
18	2230	22.5	1539	2 T48059	ABC transporter-li
19	2144.5	21.7	1355	2 T00961	hypothetical prote
20	2144.5	21.7	1514	2 T52080	multi resistance p
21	2140	21.6	1515	2 T52081	MRP-like ABC trans
22	2131.5	21.5	1490	2 T47840	multi resistance p
23	2129	21.5	1545	2 T46645	sulfonylurea recep
24	2123	21.4	1153	2 T26883	hypothetical prote
25	2109	21.3	1545	2 T42751	sulfonylurea recep
26	2104.5	21.3	1511	2 T42711	sulfonylurea recep
27	2098	21.2	1546	2 T42728	sulfonylurea recep
28	2069.5	20.9	1389	2 T47796	ABC transporter-li
29	2062.5	20.8	1661	2 S64800	probable membrane

## ALIGNMENTS

### RESULT 1

DVHUAR

multidrug resistance protein (cell line H69AR) - human  
N/Alternate names: multidrug resistance-associated protein (MRP)

C/Species: Homo sapiens (man)

C/Date: 30-Jun-1993 #sequence\_revision 05-Dec-1998 #text\_change 19-Jan-2001

C/Accession: A44231; A37495

R/Cole, S.P.C.; Bhargava, G.; Gerlach, J.H.; Mackie, J.E.; Grant, C.E.; Almquist, K.C.;

Science 258, 1650-1654, 1992

A/Title: Overexpression of a transporter gene in a multidrug-resistant human lung cancer

A/Reference number: A44231; MUID:93088080; PMID:1360704

A/Accession: A44231

A/Status: nucleic acid sequence not shown

A/Molecule type: mRNA

A/Residues: MAPTSGTMSRGIDPATPTSPAFRTSSCGGLVFTSGPV' 50-1531 <CO1>

A/Cross-references: GB:L05628; NID:91835658

A/Experimental source: small cell lung carcinoma cell line H69AR

A/Note: sequence extracted from NCBI backbone (NCBIP:119851); this sequence has been con

R/Cole, S.P.C.; Deeley, R.G.

Science 260, 879, 1993

A/Title: Multidrug resistance-associated protein: sequence correction.

A/Reference number: A37495; MUID:93262415; PMID:8098549

A/Accession: A37495

A/Status: not compared with conceptual translation

A/Molecule type: mRNA

A/Residues: 1-60 <CO2>

A/Cross-references: GB:L05628; NID:91835658

A/Note: sequence extracted from NCBI backbone (NCBIP:131929)

C/Genetics:

A/Gene: GDB:MRP

A/Cross-references: GDB:136335; OMIM:158343

A/Map position: 16p13.1-16p13.1

C/Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology

C/Keywords: antibiotic resistance; ATP; duplication; nucleotide binding; P-loop; transmem

F/661-844/Domain: ATP-binding cassette homology <ABC1>

F/678-685/Region: nucleotide-binding motif A (P-loop)

F/788-792/Region: nucleotide-binding motif B

F/130-1503/Domain: ATP-binding cassette homology <ABC2>

F/1327-1334/Region: nucleotide-binding motif A (P-loop)

F/1450-1454/Region: nucleotide-binding motif B

Query Match

Best Local Similarity 79.3%; Score 7849; DB 1; Length 1531;

Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MALRGFCSADGSDPLMDWNVTNTSNPDFTKCFQNTVLVWVPCFYLWACFPFFYLYLSRH 60

Db 1 MALRGFCSADGSDPLMDWNVTNTSNPDFTKCFQNTVLVWVPCFYLWACFPFFYLYLSRH 60

QY 61 DRGIQMTPLNKTALGFLWVCWADLFVSFMRSGIFLAPVFLVSPILLGITTLA 120

Db 61 DRGIQMTPLNKTALGFLWVCWADLFVSFMRSGIFLAPVFLVSPILLGITTLA 120

QY 121 TFLIQLERRKGVSSGIMLTFWLVALVACALAILRSKIMTALKEDAQVDLFRDITFYVYFS 180  
DB 121 TFLIQLERRKGVSSGIMLTFWLVALVACALAILRSKIMTALKEDAQVDLFRDITFYVYFS 180  
QY 181 LLLIQLVLSCFSDRSPLFSETIHDNPNCPSESSASFLSRITFWWIITGLIVRGVQPLEGSD 240  
DB 181 LLLIQLVLSCFSDRSPLFSETIHDNPNCPSESSASFLSRITFWWIITGLIVRGVQPLEGSD 240  
QY 241 LMSLNKEDTSEQVVPVLVQNMKECAKTRKQPVVYSSKDPAPKPKSSKVDANEVEAL 300  
DB 241 LMSLNKEDTSEQVVPVLVQNMKECAKTRKQPVVYSSKDPAPKPKSSKVDANEVEAL 300  
QY 301 IVKSPKQKWNPSLFVLYKTFPGPYFLMSFFFAKIHDLMMFSGPQILKLLIKFVNUTKAPD 360  
DB 301 IVKSPKQKWNPSLFVLYKTFPGPYFLMSFFFAKIHDLMMFSGPQILKLLIKFVNUTKAPD 360  
QY 361 WQGYFYTVLLFVTACILQTLVLHQYFHCIFVSGMRKTAIVIGAVYRKALVITNSARKSSTV 420  
DB 361 WQGYFYTVLLFVTACILQTLVLHQYFHCIFVSGMRKTAIVIGAVYRKALVITNSARKSSTV 420  
QY 421 GEIVNLSVDAQRFMDLATYINNIWSAPLOVILALYLLMLNLPSPVLAGVAVMLMVPVN 480  
DB 421 GEIVNLSVDAQRFMDLATYINNIWSAPLOVILALYLLMLNLPSPVLAGVAVMLMVPVN 480  
QY 481 AVAMKTKTYQVAHMKSKDNRIKLMEILNGIKVLKYAWELAFKDKVLAIROBELKVLK 540  
DB 481 AVAMKTKTYQVAHMKSKDNRIKLMEILNGIKVLKYAWELAFKDKVLAIROBELKVLK 540  
QY 541 KSAYLSAVGTFTWCTPPLVALCTFAVYVTTIDENNILDAQTAFAVSLALFNILRFPNLILP 600  
DB 541 KSAYLSAVGTFTWCTPPLVALCTFAVYVTTIDENNILDAQTAFAVSLALFNILRFPNLILP 600  
QY 601 WYISSIVQASVSLKBLRFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWARSDDPT 660  
DB 601 WYISSIVQASVSLKBLRFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWARSDDPT 660  
QY 661 LNGITFTSPIEGALVAVGVGGGKSLLSALLAEMDKVEGHVAIKGSVAVYVPOQAWIOND 720  
DB 661 LNGITFTSPIEGALVAVGVGGGKSLLSALLAEMDKVEGHVAIKGSVAVYVPOQAWIOND 720  
QY 721 SURENILFGCOLLEPYRSVIOACALLPDLPLPSGDRTEIGEKNVLSGGOKQKRVSLAR 780  
DB 721 SURENILFGCOLLEPYRSVIOACALLPDLPLPSGDRTEIGEKNVLSGGOKQKRVSLAR 780  
QY 781 AVYSNADILYLPDDPLSADVAHVGRKH FENVIGPKMLKKNKTRILVTHSMSYLPQVDVILV 840  
DB 781 AVYSNADILYLPDDPLSADVAHVGRKH FENVIGPKMLKKNKTRILVTHSMSYLPQVDVILV 840  
QY 841 MSGGKISEMGYSQELLARDGAFAPLRTYASTEQQDAEENGVTGVSFGPKKEAKOMENG 900  
DB 841 MSGGKISEMGYSQELLARDGAFAPLRTYASTEQQDAEENGVTGVSFGPKKEAKOMENG 900  
QY 901 LVTDAGLQOLQOLSSSSSYSGDISRHNSHTAELQKAEAKKEETWKLMEADKAQTQVKL 960  
DB 901 LVTDAGLQOLQOLSSSSSYSGDISRHNSHTAELQKAEAKKEETWKLMEADKAQTQVKL 960  
QY 961 SVYDYMKAIGLFIISFLSIFLFCMCHVSAASNYWLSLWTDPIVNGTQEHTKVRLSVYG 1020  
DB 961 SVYDYMKAIGLFIISFLSIFLFCMCHVSAASNYWLSLWTDPIVNGTQEHTKVRLSVYG 1020  
QY 1021 ALGISOGIAVFGYSMAVSIIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVRFSKEL 1080  
DB 1021 ALGISOGIAVFGYSMAVSIIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVRFSKEL 1080  
QY 1081 DTVDNSIMEVIMFMGSLFNIVIGACIVILLATPIAAIIIPBLGLIYFPVFYVASSRQL 1140  
DB 1081 DTVDNSIMEVIMFMGSLFNIVIGACIVILLATPIAAIIIPBLGLIYFPVFYVASSRQL 1140  
QY 1141 KRLESVSRSPVYSHNETLLGVSVIRAFEEQERFTHQSDLKVDENOKAYYPSIVANRWLA 1200  
DB 1141 KRLESVSRSPVYSHNETLLGVSVIRAFEEQERFTHQSDLKVDENOKAYYPSIVANRWLA 1200

QY 1201 VRLECVGNCIVLFAALFAVISRHSLSAGLVGLSVSYSLOVTTYLNLWLRMSSEMETNIVA 1260  
DB 1201 VRLECVGNCIVLFAALFAVISRHSLSAGLVGLSVSYSLOVTTYLNLWLRMSSEMETNIVA 1260  
QY 1261 VERLKEYSETKEAPWQIQETAPPSSWPQGVFRPNRYCLAYREDLDFVLRHINVTINGG 1320  
DB 1261 VERLKEYSETKEAPWQIQETAPPSSWPQGVFRPNRYCLAYREDLDFVLRHINVTINGG 1320  
QY 1321 EKVGIIVGRTGAGKSSLTGLFRINESAEGEIIDGINIAKILGLHDLRFKTIIPQDPVLF 1380  
DB 1321 EKVGIIVGRTGAGKSSLTGLFRINESAEGEIIDGINIAKILGLHDLRFKTIIPQDPVLF 1380  
QY 1381 SGLSRMLNDPFSQYSDREVWTSLELAHLKDFVSALPDKLDHECAGGGENLSVGQRLVCL 1440  
DB 1381 SGLSRMLNDPFSQYSDREVWTSLELAHLKDFVSALPDKLDHECAGGGENLSVGQRLVCL 1440  
QY 1441 ARALLRKTKIILVDEATAAVALDLEDDLIQSTIRITQFEDCTVLTIAHRLNTTMDYTRVIL 1500  
DB 1441 ARALLRKTKIILVDEATAAVALDLEDDLIQSTIRITQFEDCTVLTIAHRLNTTMDYTRVIL 1500  
QY 1501 DKGEIOEYGAPSDILLQORGLFYSMAXAGLV 1531  
DB 1501 DKGEIOEYGAPSDILLQORGLFYSMAXAGLV 1531

RESULT 2  
JE0336  
canalicular multispecific organic anion transporter - human  
C:Species: Homo sapiens (man)  
C:Date: 05-Feb-1999 #sequence\_revision 05-Feb-1999 #text\_change 09-Jul-2004  
C:Accession: JE0336  
R:Uchiumi, T.; Hinoshita, E.; Haga, S.; Nakamura, T.; Tanaka, T.; Toh, S.; Furukawa, M.;  
Biochem. Biophys. Res. Commun. 252, 103-110, 1998  
A:Title: Isolation of a novel human canalicular multispecific organic anion transporter,  
t.  
A:Reference number: JE0336; MUID:99032812; PMID:9813153  
A:Accession: JE0336  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1527 <UCH>  
A:Cross-references: UNIPROT:O15438; GB:AF083552  
C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology  
F,1306-1499/Domain: ATP-binding cassette homology <ABC2>

Query Match 45.3%; Score 4484.5; DB 2; Length 1527;  
Best Local Similarity 56.6%; Pred. No. 1.1e-294;  
Matches 871; Conservative 261; Mismatches 372; Indels 35; Gaps 9;

QY 8 SADGSDPLWDMNVTWNTSNPDFTKCFQNTVLVWVPCFVLWACFPFYFLYLSRHRDGVYIQM 67  
DB 7 SGEIGSKFWDNSLVHTENPDLTPTCFQNSLLWVPCYILWVALPCYLLYLRRHRCGYILL 66  
QY 68 TPLNKTALGFLWIVCMADLFYSFWERSRGIFLAPVFLVSPTLTGITTLATLFLIQLE 127  
DB 67 SHLSKLMWLVGLWCVSMADLFVSPHGLVHGRAPAFVFFVPLVGVVTMLLATLLIQYE 126  
QY 128 RRKGVQSGMLTFWLVALVCAALILRSKIMTALKEDAQVDLFRDITFYVFSLLLQLV 187  
DB 127 RLQGVQSGVLLIIPWFLCVCAIIPFRSKILLAKAEISDPFRFTFYIHFAVLVLSALI 186  
QY 188 LSCFSDSPFLFSETIHDNPNCPSESSASFLSRITFWWIITGLIVRGVQPLEGSDLSLNKE 247  
DB 187 LACFREKPPFSAKNVDNPNPETSAGLSRLFPWWTMAIYGRHPLLEKDLWSLKEE 246  
QY 248 DTSQVYVPLVKNWKECAKTRKQPVVYSSKDPAPQPKSSKVDANEVEALIVKSPQK 307  
DB 247 DRSQMVVQQLLEAWRKQEKQTAARH-----KASAPGK---NAGGEVLLGARPRP 294  
QY 308 EWNPSLFKVLKTFPGPYFLMSFFFAKIHDLMMFSGPQILKLLIKFVNUTKAPDQGVYFT 367  
DB 295 R-KPSFUKALLATFGSSFLISACFKLIQDLSLFINPOLLSILIRFISNPNMPSGWMGFLVA 353





```

Db      344  GFKSSNSYVWFYICAILMFVAVTLIQSFCLQSYQHCFVLGKVRVTTVMSSYIKKALTJ 403
Qy      411  TNSARKSSVGVIVNLSYDAQRFMDLATYINNIWSAPLOVTLALYLWNLNPGSVLAGV 470
Db      404  SNLARKQYITGETVNLMSVDSQKMDATNMQVLWSSVIQITLSIFFLRELPGSILAGV 463
Qy      471  AYVNLMPVNAVMAKTKTYQVAHMKSKDNRIKLAMEILNGIKVLKYAWELAFKDKVLA 530
Db      464  GWVLLIPVNGVLATKIRNIQVNMKNKDKRLKIMNEILSGIKILKYPAWEPFQEQVG 523
Qy      531  IQEELKVLKKSAYLSAVGTFTWCTPFLVALCTFAVYVTTIDENNILDAQTAFLVALPN 590
Db      524  IRKELKNLRFQOLSLIFILQITPILVSVTFVSVYVLDVDSANVNAEKAFSTITLNF 583
Qy      591  ILRFPNLIPMWISSTVOASVSLKRLIFLSHEELEPDSIERPVPKGGGTNSITVRNAT 650
Db      584  ILRFPNLPMPWTSSILQASVSDRLERYLGGDDLDTSAIRVSNPD----KAVFSSAS 639
Qy      651  FTWARSDDPTLNGITFSIPEGALVAVGVQGGKSSLSALLAEKMDKVEGHVAIKGSVAY 710
Db      640  FTWDPLEATIQVNLDIKPGQLVAVGVGVGSKSLVSAMLGEMENVHGHITIOGSTAY 699
Qy      711  VPQAWIQNDSURENIFCQLEEPYRSVIOACALLPDLLEILPGSDRTEIGEKNVLSG 770
Db      700  VPQSWIQNGTIKDNILFGSEYNEKYQVVKACALLPDLLEILPGDMAEIGEKGINLSG 759
Qy      771  GOKQRVSLARAVYNADIVLFDPLSADVAHVGHIFENVIGPKMLKMKTRILYTHSMS 830
Db      760  GOKQRVSLARAAQADIVLFDPLSADVAHVGHIFENVIGPKMLKMKTRILYTHGIH 819
Qy      831  YLPQVDVIVMSGGKISEMGYSQELIARDGAFELRTYAS-TEOQDAEENGVTGVSQP 889
Db      820  FLPOVDIVVLGKGTILEKSYRDLLDKGVFARNWKTFFMKHSGPEGEATVNN----- 872
Qy      890  GKEAKOMENGLVT-----DSAG-----KOLQRLSSSSSYSGDISPHENSTABLQ-- 935
Db      873  DSEAEDDDGLIPTWEEIPEDAASLAWRRENSLRRTLSRSSSSRRRGKSLNKLKNV 932
Qy      936  ---KAEAKKEETKMEADKATGOVKLSVYDYMKAGLFTSLFIFCMCHVYSALAS 992
Db      933  NVLKEKEVEQOKLTKKEFEVTGKVFYSYIKYLOAVGWWSILFILPLGULNVAFTGS 992
Qy      993  NYWLSLWT--DDPIVNGT---QEHKTVRLSVYALGATISQIAVFGYSMAVSGGILASRL 1048
Db      993  NLWLSAWTSDNLNGTNNSSSHRDMRIGVFGALGLAQICLLISTLSIYACRNASKAL 1052
Qy      1049  HVDLLHSILRSPMSFFPTPSGNLVRFSKELDTVDMSMIPVIMKPMGSLFNVIACIVI 1108
Db      1053  HQQLLTNILRAPMRFPTPTGRIVNRFSGDISTVDDLLPQTLRSWMMCFFGIAGTLVMI 1112
Qy      1109  LLATPIAAIILPLGLIYFFQRFYVASSROLKRLSVRSRSPVYSHFNETLLGVSVIRAF 1168
Db      1113  CMATPFAIILPLSLYTSQVYFVATSRQRLRSDSVTKSPIYSHFSTVTVGLPIRAF 1172
Qy      1169  EQEERFIHOSDLKVDNOKAYPSIVANRWLVRLECVGNCIVLFAALFAVISRHSLSAG 1228
Db      1173  EQQRFPLANNEKQIDINQKCVSWITSNRLAIRLELUGNLVVFCSALLVIYRKTLDG 1232
Qy      1229  LVGLSVSYLQVTTYLNWLVRMSSEMETNIVAVERLKEYSETEKAPWQIOBTAPSSWP 1288
Db      1233  VVGFVLSNALNITQTLNWLVRMTSEATNIVAVERISEYINVENEAPW-VTDKRPADWP 1291
Qy      1289  QVGRVEFRNYCLRYREDLDFVLRHINVTNGEKGIVGRTGAGKSSLTGLFRINESAE 1348
Db      1292  RHGEIQFNQYQVRYPELDELKGTICNTIKSGEKVGWGTGAGKSSLNCLFRILESAG 1351
Qy      1349  GEIIDIINIAKIGLHDLRFKTIIPQDPVLFSGSLRMNLDPFQSYSDSEVWTSLELAHL 1408
Db      1352  GQIIDGIDVASIGLHDLRERLTIIPQDPILFSGSLRMNLDPFNKYSDEVRWALELAHL 1411
Qy      1409  KDFVSLPDKLDHECAEGENISVGQRQVLVCLARALLRKTILVLDEATAVLDLEDDLI 1468

```

```

Db      1412  RSFVSGLOGLLSEVTEGGDNLSTGQROLLCIGRAVLRRKSKILVLDENTAAVDLETSII 1471
Qy      1469  OSTRTQFQEDCTVLTIAHRLNTIMDYTRVILVDKGEIOEYCAPSLLQOQGLFYMAKDA 1528
Db      1472  QTTIRKFSQCTVITIAHRLTIMDSKIMWLDNGKIVYSGPEELLNRRGFSYLMAXEA 1531
Qy      1529  GL 1530
Db      1532  GI 1533

RESULT 5
E89447
protein F57C12.4 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: E89447
R:Anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolo-
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_el-
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; an-
A:Accession: E89447
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1494 <STO>
A:Cross-references: UNIPROT:Q20943; GB:chr_X; PIDN:AAA83299.1; PID:g1118071; GSPDB:GN060
C:Genetics:
A:Map position: X
C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology

Query Match      33.2%; Score 3289; DB 2; Length 1494;
Best Local Similarity 47.6%; Pred. No. 9,7e-214;
Matches 696; Conservative 239; Mismatches 435; Indels 92; Gaps 18;

Qy      119  LATLIQIERRRGVSSGIMLTFWLVALCALAILRSLKIMTALKEDAQVDLFRDITFYVY 178
Db      76  LALILTACKNKGITSGVITLYLLVVVGIGPEPRFVLSGFIYNEALEGIRATLYIIA 135
Qy      179  FSLLIQLVLCFSRDSPLFSETHIDPNPCPESSASFLSRITFWITGLIVRGYQPLRG 238
Db      136  FTFSALELFLCCFAD---VPSDMYKSESCPEYATSFNRLTFOWFTGLAYLGNKSLN 192
Qy      239  SDSLNLKEDSEQVVLVKNWKECAKTRQPKVYVSSKDP--AOPKSSKVDANEE 296
Db      193  EDLMDLNEIDKAENLIPSMQNLKPRIDRYH-QNFK-----KPSAALPKN----- 237
Qy      297  VEALIVKSPQKEWNPSPFKVLYKTFGPYFLMSFFFKAIHDLMMFGSPQILKLLIKFVNDT 356
Db      238  -----HPSFVIPFTKYKTYLLAGFFYKLCFDMQLQFLAPOLLKQLIGFIEDK 284
Qy      357  KAPDQGVFYTVLLFVTAQLQTLVHQYFHICFVSGMIKTAIVGAVRKALVITNSARK 416
Db      285  NOPVWIGSGIVGIMFSSFLQSMFLHQYVSHFRMGVRSVLTSAVYSKALNSNEARK 344
Qy      417  SSTGEIVNLMSVDAQRFMDLATYINNIWSAPLOVTLALYLWNLNPGSVLAGVAVMLM 476
Db      345  GKTIGAIYNLMSVDLQIKQDMAPTMLFWSADPLQIFLSIYFLWKFLGVAALAGLVIIA 404
Qy      477  VPVNAVMAKTKTYQVAHMKSKDNRIKLAMEILNGIKVLKYAWELAFKDKVLAIRQEEL 536
Db      405  LPVNGLIAIQMKCKQTEQMKLKDERIKMSEILNGMKVLKYSWERSMENMVLKIREREL 464
Qy      537  KVLKKSAYLSAVGTFTWCTPFLVALCTFAVYVTTID-ENNILDAQTAFAVSLAFILRPP 595
Db      465  HILKLSYFMAAIVFESWICAPFLASVIFVYVYLDPENNVLTPEITFVALSFLDLRMP 524
Qy      596  LNILPMWISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKGGGTNSITVRNATFWAR 655
Db      525  LANVAVYGEAVQCVSVNTRLKEFFAAEMSPQT-----SISHGETDSDALEVENGLFSWSS 580

```





887 -RPFSDAAPVKGS-----TSEAQMEPSLDDVEVTGLTAGEDSVQYGRVKSATYLSYLA 940  
 970 IGLTSLFSLFLPMCNHVSALASNTYLSLWTDTPDINGTOETHKVRSLSYGALGTSQGT 1029  
 941 VGTPLCTYTLFLFLCQVASFQGYWLSUWADPVVDGQKMSALRGSIFGLLGLCQIG 1000  
 1030 VFGYSMAVIGGILASRCILHVDLLSHLSILRSPMSFFERTPSGNLVRFSKELDTVDSMIPE 1089  
 1001 LPSMAAVPLGARGASCLLFRSLLDVAVSPGFFERTPVGNLLNRFSKETDIDVDIPD 1060  
 1090 VIKPMGSLFNIVIGACTIVILLATPIAIIIPGLIYFFVRFYVASSRQLKRLSVGRS 1149  
 1061 KMTLLTYAFGLLEVLGLAVSMATPLAIVAILPLMLLYAGFQSLYVATCCQLRLSSASYS 1120  
 1150 PVSHEFNETLLGVSVTRAPEEOERFIHQSDLVADENOKAYPSIVANRWLAVRLECVGNC 1209  
 1121 SVCSHLAETFGSQVVRAPQAGPTAQHDALMDENQRISFRLVADRWLANLELLGNG 1180  
 1210 IYVLAALFAVIRHRSLSAGLVLSYSLSQVTTYLNVLVRMSSEMETNIVAVERLKEYSE 1269  
 1181 LVEVAATCAVLKSAHLSAGLAGFSYSAALQVTTQLQWVRSWTDLENSVAVERQDVH 1240  
 1270 TEKEAPWQIETAPSSHPQVRVFNRYNCLRYREDLDFVLRHINVTINGGKVGIVGT 1329  
 1241 TPKEAPWRLPSSAAQPLPFCGQIEFRDFGLRHPPELPMVAGVSLKIHAGEKVGIVGT 1300  
 1330 GAGKSLTLGLFRINESAGEIILIDGINIAKIGLHDLRFKITIIPQDPVLFSGSLRMNLD 1389  
 1301 GAGKSLTWGLRLQRETEGGIWDGVPITDGLTLRSRTIIPQDPVLPFGSLRMNLD 1360  
 1390 PFSQSDVEEWTSLRLAHKDFVSALPKDLHECAEGGENLSVGQRLVCLARALLRKT 1449  
 1361 LIQENTDEGIWALETVOLKAFVTSPLPGLQVCEGSGQDLSVGQKQLCLARALLRKTQ 1420  
 1450 IIVLEATAAVDLEDDLIQSTIRIQFREDCTVLTIAHLNIMDYTRVIVLDKGBIQBYG 1509  
 1421 IIVLEATAVDPDGTETIQWALERWFACTVLLIAHLRLSVNMCARVLVMDQGVASBG 1480  
 1510 APSDLLQORGLFYSMADAGL 1530  
 1481 SPAQLLAQKGFYRLAQESGL 1501

RESULT 8  
 S51863  
 N:Alternate names: protein YCF1 - yeast (Saccharomyces cerevisiae)  
 C:Species: Saccharomyces cerevisiae  
 C:Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text\_change 09-Jul-2004  
 C:Accession: S51863; A55352; S50233  
 R:Oliver, K.; Harris, D.  
 submitted to the EMBL Data Library, February 1995  
 A:Reference number: S51863  
 A:Accession: S51863  
 A:Molecule type: DNA  
 A:Residues: 1-1515 <OLI>  
 A:Cross-references: UNIPROT:P39109; EMBL:Z48179; NID:G665657; PIDN:CAA88217.1; PID:G6656  
 J. Biol. Chem. 269, 22853-22857, 1994  
 A:Title: A yeast metal resistance protein similar to human cystic fibrosis transmembrane  
 A:Reference number: A55352; MUID:94357936; PMID:7521334  
 A:Accession: A55352  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: DNA  
 A:Residues: 1-679, 'R', 681-1515 <SZC>  
 A:Cross-references: GB:L35237; NID:G556464; PIDN:AAA50353.1; PID:G556465  
 C:Genetics:  
 A:Gene: SGD:YCF1; MIPS:YDR135C  
 A:Cross-references: SGD:S0002542; MIPS:YDR135C.  
 A:Map position: 4R  
 C:Function:  
 A:Description: required for cadmium resistance  
 C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology

C:Keywords: ATP; nucleotide binding; P-loop; transmembrane protein; yeast vacuole  
 F:287-308/Domain: transmembrane #status predicted <TM1>  
 F:345-366/Domain: transmembrane #status predicted <TM2>  
 F:421-442/Domain: transmembrane #status predicted <TM3>  
 F:446-467/Domain: transmembrane #status predicted <TM4>  
 F:534-555/Domain: transmembrane #status predicted <TM5>  
 F:558-580/Domain: transmembrane #status predicted <TM6>  
 F:646-829/Domain: ATP-binding cassette homology <ABC1>  
 F:663-670/Region: nucleotide-binding motif A (P-loop)  
 F:951-972/Domain: transmembrane #status predicted <TM7>  
 F:995-1016/Domain: transmembrane #status predicted <TM8>  
 F:1068-1088/Domain: transmembrane #status predicted <TM9>  
 F:1092-1113/Domain: transmembrane #status predicted <TM10>  
 F:1179-1200/Domain: transmembrane #status predicted <TM11>  
 F:1208-1229/Domain: transmembrane #status predicted <TM12>  
 F:1289-1483/Domain: ATP-binding cassette homology <ABC2>  
 F:1306-1313/Region: nucleotide-binding motif A (P-loop)

Query Match 28.7%; Score 2838; DB 1; Length 1515;  
 Best Local Similarity 40.3%; Pred. No. 3.4e-183;  
 Matches 632; Conservative 295; Mismatches 495; Indels 146; Gaps 27;  
 QY 28 DFTKCFQNTVLVWVPCFYLWACFPFYF-----LYLSRHD-----RGYIQMTPLNKT 75  
 DB 27 DFTQCFDGVIL-----NLSAIFMITTIGIRDVLNLCCKHSGIKYRRNWIIVSRMALVLL 81  
 QY 76 ALGFLMLIVCWADILFYSEWERSRGIFLAPVLPVSPFTLLGITLLATFLIQLERRKGVSS 135  
 DB 82 BIAP-----VSLASNLISKEEAEEN---FTIVSQYASTMLSLFVALALHWEYDR--SVVAN 132  
 QY 136 GIMLTFLVALVALCALAILRSKIMTALKEDAQVDLFRDITFVYVFS-----LLLQLVLSC 190  
 DB 133 TVLIFYMLFETFGNFAKILNLIIRHYEG-----IWSGGTGFTLTFQVITC 180  
 QY 191 PS-----DRSPL-----FSETHDNPCEPESASFLSRITFWITGLIVRGYRQPLEG 238  
 DB 181 ASILLLEALPKPLMPHQHIQTILTRKPNPYDSANIFSRITFSWMSGLMKTGYEKYLVE 240  
 QY 239 SDLWSLNKEDTSEOVVPLVKNWKECAKTRKQPVVYSSKDPAPQKSESSKVDANEVE 298  
 DB 241 ADLYLPRNFSSEELSOKLEKNWENEL-----KQKS----- 271  
 QY 299 ALIVKSPQKWNPSLFVLYKTFGPFYLMGFFFKAIHDLMMFSGFQILKLIKFNVD--- 355  
 DB 272 -----NPSLSWAICRTFGSKMLAAFAFKAIHDVLAFTQPLRLILIKFVTDVNS 320  
 QY 356 -----TKAPDWQGYFTYVLLFVTAQLTLVHLQYPHICFVSGMRKTA 398  
 DB 321 ERQDHSLOGFENNHPQKLPIVRGFLIAFAMFLVGTQTSVLHQYFLNVTGMTYKSA 380  
 QY 399 VIGAVYRKALVITNSARKSSTVGEIVNLSMVDQRFMDLATYINMWSAPLQVILALYLL 458  
 DB 381 LTALIYQKSLVLSNEASGLSTGDIVNLSMVDVQKLDQLTQWNLNLSGPPQIILCYSL 440  
 QY 459 WNLGPSVLGAVVWVLMVPMVNAVMAKTKTYQVAHMKSKDNRIKMLNEILINGIKVLKLY 518  
 DB 441 YKLIGNSMWVGVIILVIMPLNSPLMRIOKKLQKSKYKXDRTRVISEILNLSKELKY 500  
 QY 519 AWELAFKDKVLAIR-QBELKVLKKSAYLSAVGTFTVCTPFLVALCTFAVYVTTDENNIL 577  
 DB 501 AWEKPYREKLEEVNKNELKNLTKLGCYMAVTSQFQNIIVPLVSCCTFAVFPV-VTEDEAL 559  
 QY 578 DAQTAFFVSLAFNLIRFPLNLIPMVISSIVQASVSLKRLRIFLISHEELEPDSIRRRPVKD 637  
 DB 560 TTDLVFFALFLNLLSPFLMIIPNLNSPIEASVISIGRLFTFTNEELQDPSVORLPKVK 619  
 QY 638 GGGTNSITV-RNATFTTWARSD--PPTLNGITFTSIPEGALVAVVGQVCGKSSLLSALLAE 694  
 DB 620 NIGDVAINIGDGTATFLWQRPVKVVALKNINFOAKKGNLTCIVGVKSGKTKALLSCMLGD 679  
 QY 695 MDKVEGHVAKGSAVYVPOQAQWONDSLRNILLFGCQLBEPYPSYRVSVOACALLPDLTLP 754  
 DB 680 LFRVKGFATVHGSVAVYVSVQVFWIMNGVTKENILLFGHRYDAEFYEKTIKACALTIDLAILM 739

755 SGRTEIGKGVNLSGGQKQRYSLARAVYSNADIYLFDDPLSAVDHVGKHFENVIGPK 814  
 740 DGDGKTIVGKGLSGLGGQKARLSARAVARADTYLLDDPLAAVDEHVARHLIEHVLGN 799  
 815 GMLKNKTRILVTHSMYSLPOVDVIVMSGGKISEMGSOELLA-RDGPAPBFLRYASTE 873  
 800 GLHTKTKVLTATKVSALSADSIALLDNGEITQOQTDEITKADSPMLKLLANNYK-- 857  
 874 QBDQAEENGVTGSGPGKEAKQWENGMLVTDGAGQORQLSSSSSSSGDISRHNSTAE 933  
 858 -----XNNGKSNFPGDSSSESSVRESIPV-EGEELQOLKLNLDLPGNSDAISLRASDAT 911  
 934 LOKAEAKKEETWKLMEADKAQTKVLSVYNDYMK-----IGLFIPLSIFLFCNHYVS 988  
 912 IGSIDFGDD--NIAKREHREGQKVKWNIYEVAKACNPKSCVCFILFIVISMFL----- 964  
 989 ALASNYWLSLWTDVFNQGTQHTKVLRSVYALGISQGIAGVGYSMASVI--GGILASRC 1047  
 965 SYMGNVWLKHWSVNSRYSGNSPNAAYLAIFYALGIGSALATL1QITVLMVFCTIHASKY 1024  
 1048 LHVLDLHSLRSGPMGFERTPGNLVNRFSKELDTVDSMIPEVIMFGSLFNVIACIV 1107  
 1025 LNLWNTSVLRAPMTFFETPIGRILNRFSNDIYKVDALLGRTFSQFFVNAKVTFITIV 1084  
 1108 ILLATPIAAIIIPPLGLIYFFQRFVASSRQLKLESVRSRPNVYSHNETLLGVSVIRA 1167  
 1085 ICATTWQFTFIILPSVFIYQYVYRTSRRLRLDSITRSPYISHFOETLGLLATVRG 1144  
 1168 FEOERFTHQSDLVKDNOKAYPSIVANRWLAIVLEVCVGCIVLPAALFAY--LSRHS 1225  
 1145 YSQQRFSHINOCRIDNNNSAFYPSINANRWLAIVLEIGSIILGAATLSVFLKQGT 1204  
 1226 SAGLVLSYSYSLQVTVYTNLWLVRSSEMETNIVAVERLKEYSETEKAPMOIQTAPPS 1285  
 1205 TAGMGLSLSYALQITQTLNWRVMTVEVETNIVSVERIKEYADLKSEAPLIVEGHRPPK 1264  
 1286 SWPQVGRVFRNYCLRYREDLDFVLRHNVITNGGKGVIGRTGAGKSSLLTGLFRINE 1345  
 1265 EWPSQGDIFKNYSTRYRPELDLVLRHNIHKNPKVIGVGRGAGKSSLLTALFRMIE 1324  
 1346 SAGHIIIDGINIAKGLHDLRPFKTIIPQDPVLSGSLRMLNDPFSQYSDDEVWTSLEL 1405  
 1325 ASEGNIVIDNIAINEIGLYDLRHKLSTIIPQDSQVFEGTVENIDPINQYTDIAIWRLEL 1384  
 1406 AHLKDFVSALP--DKLDHECAEGENLSVGORQLVCLARALLRKTILVLDEATAVDLET 1464  
 1385 SHLKEHVLMSNDGLDAQTEGGGNLSVGORQLCLARMLVPSKILVLDEATAVDVET 1444  
 1465 DDLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVLDKGEIOEYGA PSDLL--QORGLFYS 1523  
 1445 DKVQVETIRTAFKDRTILTIAHRLNTIMDSRIIVLDNGKVAEPDPSGQLLSDNKSIFY 1504  
 1524 MAKDAGLV 1531  
 1505 LCWEAGLV 1512

RESULT 9  
 T20434  
 Hypothetical protein E03G2.2 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
 C:Accession: T20434  
 R:McMurray, A.  
 submitted to the EMBL Data Library, November 1995  
 A:Reference number: Z19275  
 A:Accession: T20434  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-1398 <WIL>  
 A:Cross-references: UNIPROT:Q19048; EMBL:Z68113; PIDN:CAA92148.1; GSPDB:GN00028; CESP:EQ  
 A:Experimental source: clone E03G2

C:Genetics:  
 A:Gene: CESP:E03G2.2  
 A:Map position: X  
 A:Introns: 73/3; 113/3; 176/1; 216/1; 258/2; 384/3; 470/3; 549/1; 644/3; 757/3; 764/2; 8  
 C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology

Query Match 27.6%; Score 2732; DB 2; Length 1398;  
 Best Local Similarity 41.1%; Pred. No. 4.5e-176;  
 Matches 591; Conservative 287; Mismatches 472; Indels 88; Gaps 19;

QY 115 ITTLATLTLQLERKGVQSSGIMLTFWLVALVCALAILRSKIMTALKED--AQVDLRED 172  
 DB 27 LALLIATNEV---RRAGIHSSGFLFCIWMFLFAVAAPPEFYQMMTGTSGSPDELVARIDFFRY 83  
 QY 173 ITFYVYFSLIQLVLSGFSRSLFSETIHDPN-----PCPESSASFLSRITFWWITGL 227  
 DB 84 VAYLTYFLVVAEFVLFVSDPPFM-----PRGYNLKCPEENANFISRQLLWFTQI 136  
 QY 228 IVRGVRQPLEGSDMLSNKEDTSEQVVPVLVKNWKECAKTRKQPVK-----VYSS 279  
 DB 137 ISLGYERTLVADDVPEMDSQMDQEYLKARWKTMLKQTEKAREKQVLDKRRARTGSE 196  
 QY 280 KDPQPKESSKVDAN--EEVEALIVKSPQKWNPSLFVLYKTPGYFLMSPEFFKAHDLM 338  
 DB 197 KAPLLGTNNYGANVLDKDRVIVQ-----PSVITLQIMKWEILGGSPFKLSLDL 249  
 QY 339 MFSGQIILKLIKFVNDTKAPDMQGYFYTVLLFVTAQLQTLVHLQYHFCFVSGMRIKTA 398  
 DB 250 QFANPTFLNYLLFIETPNALPLINGIGLAVGLFAGQIKSLFMNTYFIAMTRVGAKIQTM 309  
 QY 399 VIGAVRKALVITNSARKSTVGEIVNLSMDVDAQRFMDLATVINMWSAPLOVIALYLL 458  
 DB 310 LSCAVYEKSLLSNTARRERTVGMENVILSDVDRFRMITPQIQOYSSSPFQIICMVL 369  
 QY 459 WNLNPGSVLAGVAVMLVMPVNAVMAKTKYOVAAHMSKNRIKLMNEILNGIKVLKLY 518  
 DB 370 SQTIGVAVAGIVVMISIVINICVSMITKRWQLRMKYKDBIRILINEVLNGIKVKVLS 429  
 QY 519 AWEIAFKDKVLAIROEELKVLKSAVLSAVGTFTWVCTPFLVALCTFAVYVITDENNIID 578  
 DB 430 AWEIAMEETIERVRDKELKMIKQSALLKTFADCLNVGAPVFAVLSFTFVFLIDPKVLT 489  
 QY 579 AQTAFVSLALFNIILRFPNLPIMVLISSIVQASVSLKRLRIFLSHBELSPDTERRPVKG 638  
 DB 490 PNIAFVSLFNLRLGPMMAELVAQVQLVVSNNKRVRTFLCEKEVDTAADKE--IRGE 548  
 QY 639 GGTNSITVRNATFTWASDDPPTLNGITFISIEGALVAVVGVQVCGCKSSLSLALLEMDKV 598  
 DB 549 LYTNVTEHSGSFAWDSAEARILSDIEFLAGSKELVTVVSGSGKSLLLAALGEMEKV 608  
 QY 699 EGHVAIKGSVAVPQQAWIQNDSLRNLTFLGQLEPYYRSVIOACALLPDLLEILPSGDR 758  
 DB 609 CGYGVGRSVAYLSQPPWILNQSLKKNVLMQADLNDVLYKKVIESCALKEDLKQLPDGDD 668  
 QY 759 TEIGEGVNLGGQKQVRSARAVYSNADIYLFDDPLSAVDHVGKHFENVIGPKMLK 818  
 DB 669 TEIGEGVNLGGQKQVRSARAVYSNADIYLFDDPLSAVDHVGKHFENVIGPKMLK 728  
 QY 819 NKTRILVTHSMYSLPOVDVIVMSGGKISEMGSOELLA-RDGPAPBFLRYASTEQEDA 878  
 DB 729 HTTRILVTNCTSFQESGKIIVMKG-----ETRYESSG 761  
 QY 879 EENGVTGSGPGKEAKQWENGMLVTDGAGQORQLSSSSSSSGDISRHNSTAE LQKAE 938  
 DB 762 EESG-----GEENSDILPGSIAGSRMSLSR--LSK-----ISRKSKSSIVEK-- 804  
 QY 939 AKKETWKLMEADKAQTKVLSVYNDYMKIAGLFIPLSIFLFCNHYVS--ALASNYWLS 997  
 DB 805 -KKPDA--LITKEEAAIGRVNPGVYLLFYKAMGIVTYVLPVAIVLVNVSFALGRSLMT 861  
 QY 998 LWTDDPIVNGTQEHKTV--RLSVYCALGISQGIAGVGYSMASVIGGILASRC LHVLDLHS 1055  
 DB 862 AWSNDANIDINHPDNTMSVGARLGVAGFGITVIFLFFSLVLLIGGVAASKNLHKLPLHN 921



QY 1227 AGLVGLSVSYSLQVTTYLNLWLRMSSEMTNIVAVRLKYESETEKAPQIQETAPPSS 1286  
Db 1173 SGLVGLSLSYAVQIQSLTFVVRQSDVETNIVSYERMLEYIGLSEAPSIIPDRPPEG 1232  
QY 1287 WQGVGRVFRNYCLARYREDLDVLRHINTVINGGKVGIGTGAGKSLTGLGRINES 1346  
Db 1233 WPSGAIRPDHYSVRYRENLPVLVNDISVNIHQPKIGIGVGTGAGKSTLTALFRLIEP 1292  
QY 1347 ABGEIITIDGINIAKIGLHDLREKTIIPQDPVLFSGSLRMNLDPPFSQYSDBEVWTSLELA 1406  
Db 1293 TSGDITQLODDINITSGLHDLRSRLAIIPQENQAFSGTIRENLDPNANATDEEIHWALEAA 1352  
QY 1407 HLKDFVSALPKDLDEHCAEGENLSVGQRQLVCLARALLKTKIILVLEATAAVALDETDD 1466  
Db 1353 SLKQFIQTLDDGLXSVTVEGGANLSSGQRQLMCLTRALLTPRVLLEDEATAAVALDETDA 1412  
QY 1467 LIQSTIRQFEDCTVLTIAHRLNTIMDYTRVILDKGEIOEYGCAPSDILQOR-GLFYSWA 1525  
Db 1413 IVQRTIRERFNDRTILTIAHRLNTVMSNRILVLDHGKVFEPDSTKCLLENKASLIFYSLA 1472  
QY 1526 KDAGLV 1531  
Db 1473 KESGLI 1478

RESULT 11  
T01369  
ABC transporter AtMRP2 [imported] - Arabidopsis thaliana  
N:Alternate names: multidrug-resistance protein homolog T29F13.13  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 09-Jul-2004  
C:Accession: T01369; D84759  
R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul  
submitted to the EMBL Data Library, May 1998  
A:Description: Arabidopsis thaliana chromosome II BAC T29F13 genomic sequence.  
A:Reference number: Z14179  
A:Accession: T01369  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1623 <ROI>  
A:Cross-references: UNIPROT:064590; EMBL:AC003096; NID:g3132469; PID:g3132479  
A:Experimental source: cultivar Columbia  
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;  
eues, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A:Reference number: A84420; MUID:20083487; PMID:10617197  
A:Accession: D84759  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1623 <STO>  
A:Cross-references: GB:AE002093; NID:g3132479; PIDN:ACC16268.1; GSPDB:GN00139  
C:Genetics:  
A:Gene: At2g34660; T29F13.13  
A:Map position: 2  
A:Introns: 110/3; 170/2; 183/3; 236/1; 274/2; 301/2; 309/3; 328/3; 350/3; 375/3; 411/3;  
C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology  
F:1259-1452/Domain: ATP-binding cassette homology <ABC>

Query Match 24.1%; Score 2384.5; DB 2; Length 1623;  
Best Local Similarity 35.2%; Pred. No. 1.9e-152;  
Matches 561; Conservative 274; Mismatches 534; Indels 223; Gaps 30;

QY 44 FYLWACFPF-YFYLRSRDRGYIQMTPLNKTATLG-FLMWVCWADLFYSFV----- 94  
Db 5 FIEWYCKPVPNGVWTKQVANAFGAYTPCATDSFVLGISQLVLLVLC-----LYR1W1ALXD 60  
QY 95 -----ERSRGIFLA-----PVFLVSPITLIGITL-----LATFLIQLERK 130  
Db 61 HKVERFLRSRLNYNYFALLAAVATAPLFRL---IMGISVLDGFGPLPPF----- 109  
QY 131 GVQSSGIMLTFMLVALVLCALAILRSKIMTALKEDAQVLDLFRDITFYVYFS-----LL 182

Db 110 --EAFGLGVKAPAWGVMMILMETKI-----YIRELRWYRFVAVIYALVGMV 156  
QY 183 LIQVLVS-----CFSDRSPLSETIHD--- 204  
Db 157 LLNLVLSVKEYSSVVLVLYTSEVCAQVLFGLLPMHLPNLDITYPGYNPVRSETVDDIYEY 216  
QY 205 -----PNCPRESSASFLSRITFWITGLIVRGYRQPLEGSDLSLNKEDSESQVVPVLV 258  
Db 217 EEISDGOQICPEKHANIPDKIFFSWMNPMLTGLSKRPLTEKDVWYLDTDQOTETLFTSFQ 276  
QY 259 KWKKECAKTRKQPKVYVSSKDPAPKESKVDANEVEALIVKSPKEMNPSPKVLVY 318  
Db 277 HSWDKELQK-----PO-----PWLRALN 295  
QY 319 KTFPGYFLMSFFFKAIHDLMMFSGPOILKLIKFNVDTKAPDWOQGYFYTVLLFVTAQLQT 378  
Db 296 NSLGRFWMGFWKIGNDCSQFVGPLLNLQLLKSNQE--DAPWAGYIYAFSIFVGVVFCV 354  
QY 379 LVLHQYFHCIFVSGNRKIRKAVIGAYRKALVITNSARKSSTVGEIVNLMSVDAQAFMDLA 438  
Db 355 LCEAQYFQNVVRVGYRLRSALIAAVFRKSLRLTNEGRKFKOTGKITNLMTTDAESLQOTC 414  
QY 439 TYINMINSAPLOVILALYLLMLNLGSPVLGAVVAVMLVAVPNVAMAMTKTYQVAHMSK 498  
Db 415 QSLHTMSAPFRITIIALILLYQQLGVASLIGALLLVLMFPLOTVIISKKQKLTKEGLQT 474  
QY 499 DNRIKLMNEILNGIKVLKLYAWELAFKDKVLAIROEELKVLKKSAYLSAVGTFTWCVTPF 558  
Db 475 DKRIGLMEVLAAMDVTVCYAWENSFQSKQTVTRDDELSPFKSOLLGALNFIINSIPV 534  
QY 559 LVALCTFAVYVITDENNILDQAQFVSLALFNILRPLNIPMLVSISSIVQASVSKRL-- 616  
Db 535 LVTIVSGVFTLGGD--LTPARAFTSLSLFAVLRFPFLMPLNIIITQVNVANVSKRLSE 592  
QY 617 -----RIFLSHEELEPDSIERRPVKDGGGTNSIIVRNATFTW--ARSDPTLNGITFTSP 669  
Db 593 VLATEERILLNPPIEP-----GPAISIRNGYFSWDSKGRPTLSNINLDVP 640  
QY 670 EGALVAVVGVGVCGKSSLSALLAEMDKV--EGHVAIKGSVAVVPOQAQWIONDSLRENILF 728  
Db 641 LGLSLVAVVGVSTGEGKTSLSAILBELPATSDAIVTLRGSVAVVPOVSWIFNATVRDNLIF 700  
QY 729 GCQLEPYRYSVIOACALLPDLLEILPSGDRTEIGKGNVLSGGQKORVSLARAVYSNADI 788  
Db 701 GSPFDREKVERAIDVTSKHDLELPGDLEIGERGYNISGGQKORVSMARAVYSNSDV 760  
QY 789 YLFDDPLSADVAHVGHKHFENVIGPKGMLKNTKRLIVTHSMYSYLPOVDVILVMSGGKISE 848  
Db 761 YIFDDPLSALDAHVGOQVFEKCI--KRELQKTRVLVTNQLHFLSQVDRIVLVHSGTVKE 818  
QY 849 MGSYCELLARDGAFABEFLRTVASTEQEODAENGVTGVSGPKE--AKQMENGMLVTDSEA 906  
Db 819 EGTYEELSSNGPLFORLMENAGKVEEY--SENGEAEADQTAEQPVANGNTNGLQMDGSD 876  
QY 907 GQKLQRLSSSSSYSGDISRHHSNTEAELOKAEAKKEETWKLMEADKAQGTQVKLSYVWDY 966  
Db 877 DKK-----SKGNKGGKSVLIKQBERETGVVSWRVLKRY 911  
QY 967 MKAI--GLFISFLSIFLFCMNVHVSALSNYLSLWTTDDPIVNGT--QEHYKVRLS--VYGALG 1023  
Db 912 QDALGGAVVMMLLCYVLTVEFVTSSTLSEWTD----ACTPKSHGLFLFNLYIALLS 967  
QY 1024 ISQGIATVFGYSMAVSGIGGILASRCLHVDLLHLSRSPMSFFPTSGNLVNRFSKELDTV 1083  
Db 968 FGQVLVTNTNLTWLTNSSLIAAKLUHDNMLHLSILRAPSFHTNPLGRINRFAKDLGDI 1027  
QY 1084 DSMIEPIKMFMSGLFNIVGACIVILLATPIAAIIPPLGLIYFFVQRFYVASSRQLKRL 1143  
Db 1028 DRTVAVFVNMFMFGQVSQLSTVTLVIGIVSTLSLWAIMPLLVLFYGAAYLYQNTAREVKRM 1087  
QY 1144 ESVSRSPVYSHNETLLGVSVIRAFEEQBRFTHQSDLKVDENQKAYPIPSIVANRWLAVRL 1203  
Db 1144 ESVSRSPVYSHNETLLGVSVIRAFEEQBRFTHQSDLKVDENQKAYPIPSIVANRWLAVRL 1203



```

Db 1232 PSSGSIKFDVVRVRLPELPPVHLGVSLISPMKDVIGVGRGAGKSSLLNALFRIVELE 1291
Qy 1348 EGEIIDGNIAGIKGLHDLRFKTIIPQDPVLFSGSLRMNLPFSGYSDEEVTSLSLAH 1407
Db 1292 KGRILLDECDIGRFGMLMDLRKVLGIIPQAPVLFSGTVRNLDPFSEHNDADLWESLERAH 1351
Qy 1408 LKDFVSALPKLDHCEAGGEMLSVGQRLVCLARALLKTKILVLDEATAAVIDLETDL 1467
Db 1352 LKDTIRNPLGLDAEVEAGENFSGQRLSLARALLRSKILVLDEATAAVIDRTDVL 1411
Qy 1468 IQSTRTQFEDCTVLTIAHRLNTIMDYTRVIVLDKGEIQEYCAPSDLLQOQGLFYSMAD 1527
Db 1412 IQKTREEFKSCTMLIIAHLNTIIDCDKVLVLDGKGVQFSSPENLLSNGSSFSK--- 1468
Qy 1528 AGLVGGGGGMLSRKGIIPPEEYVLTRLAEDPAEPRYRTRE 1567
Db 1469 -----MVQSTGTANAELRSITLENK-----RTRE 1493

RESULT 13
E86428
probable ABC transporter [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: E86428
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: E86428
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1495 <STO>
A:Cross-references: UNIPROT:Q9C8H0; GB:AE005172; NID:g11055918; PIDN:AAG28288.1; GSPDB:C
C:Genetics:
A:Map position: 1
C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology

Query Match 23.5%; Score 2326; DB 2; Length 1495;
Best Local Similarity 34.3%; Pred. No. 1.5e-148;
Matches 541; Conservative 314; Mismatches 566; Indels 154; Gaps 28;

Qy 2 ALRGFCS--ADGSDPLMDNNTWNTSNPDFTKCFQNTVLVWVP-----CFY-LWACFP 51
Db 5 ALNWYCKPVADGF-----WEKAVDGAAGATPCADISLVMVSHFVLLGLCFYRIWIIP- 58
Qy 52 FYFLYLSRHDRGYQMTPLNKTALGFLLIWVCWAD-----LFVSFWERSRGIFLAPV 105
Db 59 -----HNTKAIYVLRKKYKNCVLG-LIACVVEPVVLRVMGSLFDMDEETDFPPF 110
Qy 106 FLVSPILLGITTLATFLIQLERRKGVQSSGIMLTFWLVALVCLAILRSKIMTALKEDA 165
Db 111 EVASLMVEAFAMFMSMLVLIGLETQYKVEFRWYVRFGLVADAVLLDLVPLKNSIN 170
Qy 166 QVDLFRDIT---PYVYFSLILLIQLVLSCFSDRSPFSEITHDP-----NPPC 209
Db 171 RTALYLFISRCQALFGILLIYIPEL--DPYFGYHIVNEPDLNVEYDALRGGEHICP 228
Qy 210 ESSASFLSRITFWITGLIVRGYRQPLEGSDLSLNKEDTSEQVPLVKNMKCEAKTR 269
Db 229 ERHASIFSRIFYFCWITPLMQLYRKYTEKDVQWLDKWDQETETLIKRFQRCWTEESRR-- 286
Qy 270 KQPVKVVYSSKDKPAQKESKVDANEVEALIVKSPQKWNPSLFKVLKTYTGPFLMSF 329
Db 287 -----PK-----PWLRLANSLGGRFWLAG 307

```

```

Qy 330 FPKATHDLMFSGPOILKLLIKFVNDTKAPDQGWYFYTVLLFVTAQLOTLVLHQYFHICF 389
Db 308 IFKIGNDLSQFVGPVILSHLLRSMQEGD--PAWGVGVYAFIIFVGVTLGLVCEAQYFQNVW 366
Qy 390 VSGMRKATVAVYRKALVITNSARKSSTVGEIYNLMSVDAQRFMDLATYINMLWSAPL 449
Db 367 RVGFRRLRTVAAIIPKSLRLTHEARKNFPASQKVTNMTITDANALQOISQQHWSAPF 426
Qy 450 QVILALYLLMLNLGPSVLAVGVAVMVMVPMAMKTKYQVAHMKSDNRKILMNEIL 509
Db 427 RIIVSMILLYQOLGVASLFGSILFLPLQTLIISKMKLTKEGQLWTDKRVGITNEIL 486
Qy 510 NSIKVLKUYAMELAPKQVLAIRQBELKVLKKSAYLSAVGTFTWCTPPLVALCTPAPVY 569
Db 487 SMDTVKCYAMEKSPESRIGIRNEELSWFRKAQLLSAFNSFILNSIPVVVTVWSFGV 546
Qy 570 TIDENNIIDAQTAQFVSLALFNILRPLNILPMVSIIVQASVSLKRL-RIFLSHELEPD 628
Db 547 LIGGD--LTPARAFSTLSLFAVLRPLNMLNLLSQVNVANVSLQRIEBELLSEBILAQ 604
Qy 629 STIERPVDKGGGTSITVRNATFTW-ARSDPPTLNGITFSIPEGALVAVGVGGCKSSL 687
Db 605 NPPLQP-----GTPAISIKNGYFSDSKTKPTLSDINLEIPVGLVALVGTGEGKTSL 659
Qy 688 LSALLAEMDKVE-GHVAIKGSVAYVPOQAWIONDSRLNLPFCOLEBEPYRSVIOACAL 746
Db 660 ISAMLGELSHAETTSWIRGSVAYVPOQSWIFNATVRENILFGSDPESERYWRAIDATAL 719
Qy 747 LPDLILPSGDETEIGEKNVLSGKOKORVSLARAVYSNADIVLPDDPLSADVAHVXHI 806
Db 720 QHDLPLPGRDLTEIGRGNISGKOKORVSNARAVYNSDVIYIPDDPLSADVAHVXHV 779
Qy 807 PENVIGPGLMKVTKRILVTHSMSYLPQVDVIVMSGGKISRMGYSQVQELLARDGAFELP 866
Db 780 FDSCH--KDELGRKTRVLTQHLPLMDKILLVSEGMIKEGTFVELSKSGLFKKLM 837
Qy 867 RYASTEQEQDABENGVTGSGPGKEAKQOMENGLVTDGAKQOLQRLSSSSSYSGDISR 926
Db 838 ENAGKMDATQEVNTN-----DENILKLGPTVTVDVS---ERNLGSTK----- 876
Qy 927 HNSTAELOKABAKKEETWKLMEADKAGTQVYKLSVYDYMKA-GLFTSFLSIFLPMCN 985
Db 877 -----QGGKRRS-VLIQKBERETGIISNVLNRYKEAVGLVWMLILLACYLAT 924
Qy 986 HVSALASNVLSLWTDTPVNGTQEHK-----VRLSVYGALGISOGIAVFGYSMAVSIG 1040
Db 925 EVLRVSSSTWLSIWD-----QSTSKNYSFGFVIVVYVALLGFGQVAVFTNSFWLITS 977
Qy 1041 GILASRCLHVDLLHLSILRSPMSFFERTPSGNLVNRFPSKELDTVDSMPIVKMFMSGLFN 1100
Db 978 SLHAARLHDAMLSSILRAPMLFFHTNPTGRVINRFSKIDGIDRNVANLMMNFMNLQW 1037
Qy 1101 VIGACIVILLATPIAAIIPPLGLIYFFVQRFVAVASSRQKLESVSRSPVYSHNETLL 1160
Db 1038 LLSTFALLTGVTISLWAIMPLLIIFAYLYQSTSRVRLDSVTRSPITVQAQGEALN 1097
Qy 1161 GVSVTRAPFEEQBRFTHQSDLKVDENOKAYPSIVANRMLAVLECVGNCIVLFAALFAVI 1220
Db 1098 GLSSIRAYKAYDRMAKINGSNDNIRFTPLANTSSNRMLTIRLETGGVMIWLTATFAVL 1157
Qy 1221 SRHSLS-----AGLVCLSVSYSLQVTVNMLNMLRMSSEMETNIVAVERLUKSEYSEKEAP 1275
Db 1158 QNGTNNQAGFASTWGLLLSYTLNITSLSGVLRQASRAENSLSNVSERVGNVIDLPEAT 1217
Qy 1276 WOIQSTAPPSSWPQVRVEFRNYCLRYREDLDFVLRHINVTINGEKKIVGRTGAGKSS 1335
Db 1218 DIENNRNPPVCGWPGSGSKFEDVHLRYRPLGLPVHLGLTFFVFSPEKVGWGTGAGKSS 1277
Qy 1336 LTLGLFRINESAEGEIIIDGINIAKIGLHDLRFKTIIPQDPVLPFGSLRMLNDPPSQYS 1395
Db 1278 MLNALFRIVEVEKGRIMIDDCDVAKGLTDRVRVLSIIPQSPVLPFGSTVRFNIDPFSEHN 1337

```

Db	530	---	KSEKPKQVKNVE--NVQLGRVKSVYQLYIKTMGIFNSAFLIPIAHTFTVMIMR	584	
Qy	993	NYWLSLWTD	-----PVGNTQE-----HTKVLRSVYGALGISOGIAVFGYSM	1035	
Db	585	SLMUSDWNSNAIKATLSVDYLNSTSVDPGVSVE	TRLIYVAGFG---GUEMLLJAL	641	
Qy	1036	A--	VSTGIGLASRCLHVDLLHSLRSPMSFFERTPSGNLVNRFSEKELDTVDSMIP	EVIK	1092
Db	642	AFTVLITGSLRASVGLHSLPIJHALLVAPISFFDTTPTGRINRLSRDL	VIDK-LQDNIR	700	
Qy	1093	MFGSLPNVIGACIVILLATPIAAIIPPLGLIYFFVQRVYVASSRQKLES	VSRSPVY	1152	
Db	701	MCTQTLINACMIILVLISISTPIFLVCAAPLILYFYFMIYYIPTSRQKLES	ANRSPIL	760	
Qy	1153	SHFNETLLGVSVIRATEEQERFIHQSDLKVDENQKAYPSIVANRMLAVLE	CGNCIVL	1212	
Db	761	STIAESTHGASSIRAFDKTERTTTALSTNVDKFAQCRLYLSHMSNRMLATE	LELLGNTCVL	820	
Qy	1213	FAALFAVISR--HSLSAGLVGLSVSYSLQVTTVLNVLNVRMSSEMTNIV	AVERLKEYSET	1270	
Db	821	FASLSATLSIKYGLTFPGMAGLSVSVALTITEVLNLCVRSV	ESNIIVSVERVNEVQKL	880	
Qy	1271	EKEAPMOIQETAP-PSSPWQGVREFRNYCLARYREDLD	DFVLRHNVTINGEKGIVGRT	1329	
Db	881	EPEAPRIEKSENEEKWPVKGIKELDGFGRYKKNLPLVLKNI	DLKIEGGERIGVIRT	940	
Qy	1330	GACKSSLTICGLFRINSAEGEIILDGINIAKIGHDLRFKITIIPQDPVL	FCSGLRMNLD	1389	
Db	941	GSKSSLTWALYRMIEGESGTIKIDVDEITIGLHQLRSKLIIPQEPV	FSGLTRPFLND	1000	
Qy	1390	PFQSYDDEEVTLSLELAHLKDFVSALPDKLDHCASGGENLSVGQRLV	CLARALLRKT	1449	
Db	1001	PFNQYSDQIWNCLEICQLKQFAQEDDKTLDRYIAEGGKMSV	GERQLLCLCRALLRGAR	1060	
Qy	1450	ILVLEATAVAVDLETDDLQSTIRTPQEDCTVLTIAHRLNTIMDYTR	IVILDKGEIQEYG	1509	
Db	1061	IVILDEATASVDVTVDGIVQAIRQHFPQSTTISIAHRLDTIVDS	DRIVLDAGRAVEFD	1120	
Qy	1510	APSDLLIQORGLFYS	1523		
Db	1121	TPSNLLNPDSLYS	1134		

RESULT 15

S64757

probable membrane protein YLL015w - yeast (Saccharomyces cerevisiae)

N;Alternate names: hypothetical protein A255; hypothetical protein L1313

C;Species: Saccharomyces cerevisiae

C;Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text\_change 09-Jul-2004

C;Accession: S64757; S64763; P50041; S63931; S70560

R;Miosga, T.; Zimmermann, F.K.

submitted to the Protein Sequence Database, May 1996

A;Reference number: S64743

A;Accession: S64757

A;Molecule type: DNA

A;Residues: 1-1559 <MIO>

A;Cross-references: UNIPROT:P14772; EMBL:Z73120; NID:gl360184; PIDN:CAA97460.1; PID:CAAG97460.1;

A;Experimental source: strain S288C

R;Goffeau, A.; Purnelle, B.

submitted to the Protein Sequence Database, May 1996

A;Reference number: S64761

A;Accession: S64763

A;Molecule type: DNA

A;Residues: 1-1559 <GOF>

A;Cross-references: EMBL:Z73120; NID:gl360184; PIDN:CAA97460.1; PID:gl360185; G

A;Experimental source: strain S288C

R;Boy-Narcotte, E.; Damak, F.; Camonis, J.; Garreau, H.; Jacquet, M.

Gene 77, 21-30, 1989

A;Title: The C-terminal part of a gene partially homologous to CDC25 gene suppressor

A;Reference number: PS0041; MUID:89306677; PMID:2545538

A;Accession: PS0041

A;Molecule type: DNA

A:Residues: 1-255 <BOYS>  
 A:Note: the authors translated the codon CAG for residue 248 as His  
 R:Purnelle, B.; Goffeau, A.  
 submitted to the EMBL Data Library, April 1996  
 A:Description: The sequence of 32 kb on the left arm of yeast chromosome XII reveals 14  
 mly and a new ABC transporter homologous to the human multidrug resistance protein.  
 A:Reference number: S69380  
 A:Accession: S69391  
 A:Molecule type: DNA  
 A:Residues: 1-1559 <PUR>  
 A:Cross-references: EMBL:X97560; NID:g1297003; PIDN:CAA66162.1; PID:g1297015  
 R:Miogga, T.; Zimmermann, F.K.  
 Yeast 12, 693-708, 1996  
 A:Title: Sequence analysis of the CEN12 region of Saccharomyces cerevisiae on a 43.7 kb  
 e conductance regulator protein CPT.  
 A:Reference number: S70557; MUID:96405918; PMID:8810043  
 A:Accession: S70560  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-1559 <MIW>  
 A:Cross-references: EMBL:X91488; NID:g1495203; PIDN:CAA62776.1; PID:g1495208  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1995  
 C:Genetics:  
 A:Gene: SGD:BPT1; MIPS:YLL015w  
 A:Cross-references: SGD:S000938  
 A:Map position: 12L  
 C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology  
 C:Keywords: ATP; nucleotide binding; P-loop; transmembrane protein  
 F:28-44/Domain: transmembrane #status predicted <TM1>  
 F:143-159/Domain: transmembrane #status predicted <TM2>  
 F:178-194/Domain: transmembrane #status predicted <TM3>  
 F:334-350/Domain: transmembrane #status predicted <TM4>  
 F:421-437/Domain: transmembrane #status predicted <TM5>  
 F:526-542/Domain: transmembrane #status predicted <TM6>  
 F:550-566/Domain: transmembrane #status predicted <TM7>  
 F:654-847/Domain: ATP-binding cassette homology <ABC1>  
 F:672-679/Region: nucleotide-binding motif A (P-loop)  
 F:974-990/Domain: transmembrane #status predicted <TM8>  
 F:1017-1033/Domain: transmembrane #status predicted <TM9>  
 F:1099-1115/Domain: transmembrane #status predicted <TM10>  
 F:1118-1134/Domain: transmembrane #status predicted <TM11>  
 F:1212-1228/Domain: transmembrane #status predicted <TM12>  
 F:1319-1529/Domain: ATP-binding cassette homology <ABC2>  
 F:1336-1343/Region: nucleotide-binding motif A (P-loop)

Query Match 23.2%; Score 2294.5; DB 1; Length 1559;  
 Best Local Similarity 35.1%; Pred. No. 2.2e-146;  
 Matches 574; Conservative 292; Mismatches 541; Indels 227; Gaps 39;

QY	23	NTSNPDFTKCFONTVLVWVPCFLW-ACFPFVFLSLRHRDGVIOMTPLNKTALGFL-	80
DB	23	NALNP-----CFISVISAWQAVFLLIGSQWLKLYKNNKVPFRKXNFFPLPSKINSRLT	78
QY	81	-LWIVCW-----ADLPYSFWERSRGIFLAPVFLVSTLGI---TTLAT	121
DB	79	HLTNVCFQSTLIICELALVSQSDRVYPTLK-KALYNLLE-----NLGISLPTQYLA	132
QY	122	FLIQLERRKGVQSSGIMLFWLVALVCAALILRSKI-----MTALKEDAQV	167
DB	133	P-----KSTFSMGNLFPYMQFQILQLFLQRYHGSNRLTVISGQTAMILEVL	185
QY	168	DLFRDITFVYSLLLIQVLSCFSDRSLFSETHIDNPCESSASPLSRITFWITGL	227
DB	186	-LNSVAIFIIY-DLCIFPEI-----NELSEYKXNGWYP-----PVHVLSYITFWMKL	233
QY	228	IVRGYRQPLEGSDWLSLNK-EDTSEQVQVFLVKNWKKCAKTRKQPKVYVSSKDPAPQ	285
DB	234	IVETIR-----NKKIDPNQLPLPPVDNLK-----	259
QY	286	KSSKVDANEEVEALIVKSPQKEWNPFLKVLKYKTPGFLMSPFFKALHDLMFSGPOI	345
DB	260	SISKEFKANWELEKWLNRN-----SLWRAIWKSFGRTISVAMLYETTSLLSVVQPOF	312

QY	346	LKLLIKFVN---DTKAPDQGYFYTVLLFVTAQLTLVLHQVHFICFVSGMRKTAFAVIGA	402
DB	313	LRIFIDGLNPETSSKYPPLNGVFIALTLFVISVWSVFLTNQFYIGFEAGLIRGSLASL	372
QY	403	YVRKALVTNSARKSSTVEIVNLSVDA---QRPMDLATYINMISAPLQVILALYLW	459
DB	373	VYQSLRUTLAEKNEKSGDILNLSVDVLRIORFENA---QTIIGAPIQIIVVLTSLY	429
QY	460	LNLGSLVAGVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVV	519
DB	430	WLLGAVVIGLVTWAIMMPINAFLSRKVKLSKTKOMYKDMRIKTIITELLNAIKSLYA	489
QY	520	MELAFKDKVLAIROB-ELUKVKKSAVLSAVGTFTWCTPFLVALCTFAVYVVIDENNILD	578
DB	490	WEEPMMARLNHRNDMELKFKRIGIVSNLIYFAMNVCPLMVTCTFGIF-SLFSDSLPS	548
QY	579	AQTAFVSLALFNILPPLNIPMTISSIVQASVSLKRLRIFLSHEELPDSLER-RPVKD	637
DB	549	PAIVFPLSLFNILNSAIYSPSMINTIETSVMERLKSFLSDEIDDSFIERIDPSAD	608
QY	638	GGGTSITVRNATFW-----ARSDPTTLNGIT-----PSIPEGALV	674
DB	609	ERALPAIENWITFLWKSKEVLTSSQSGDNLRTDESIIGSSQIALKNIDHPEAKRGDLV	668
QY	675	AVVGQVCGKSLLSALLAEMDKVEGH-----VAIKGSVAVYVQQAQIOWNSLRNI	726
DB	669	CVVGRVAGKSTFLKAILQLPCMSGSRDSTPPKLIIRSSSVAYCSQESWIMNASVRENI	728
QY	727	LFGQLEBPYRVSVOACALLPDLEILSPGDBTEIGKGVNLSGGQKQKQVSLARAVYNA	786
DB	729	LFGHKFDQDYDLTIKACQLLPDLKILPDGDETLVGEKGISLGGQKARLSARAVYNSA	788
QY	787	DIYLFDDPLSAVDAHVGHKIPENV-IGPKGMLKKNKTRILVTHSMGYLPQVDVIIVMSGK	845
DB	789	DIYLLDDILSAVDAVSNKIIIEYVLGKALKKNTIITNTVTSILKHSQMIYALENGE	848
QY	846	ISEMGSYQELLAR---DGAPAEFLRYAS-----TEQPDQAEENGVTGVSFGPKAKOME	897
DB	849	IVEQNGVEDVMNRKNTSKLKLLEFDFSPIDNGNESDVQTEHRSESEVDEF-----	900
QY	898	NGMLVTDGAGQLQOLQOLSSSSYSODISRHHNSTAELOKAEAKKSETWKL-----	947
DB	901	-----LQKLVTESETEDEVVT---ESELELIKANSRRASLALTRPRPFVGAOL	945
QY	948	-----MEADKAOTGOVKLSVYWDYMKAIGLFISFLSIFLPM-CNHVSALASNYWLSW	999
DB	946	DSVKTAQAKAEKTEGVRVTKIYIAYIKACGV-LGVVLFLEPMLITRFDLAENFWLKW	1004
QY	1000	TDDPIVNGTOEHTKVLRSVYGALGISOGIAVFGYSMAVSI---GGILASRCLHVDLLHSI	1056
DB	1005	SESNEKNGSNERVMFVGVYSLIGVAS--AAFNNLRSIMMLLYCSIRGSKLHESMAKSV	1062
QY	1057	LRSFMSFPTSPGNLVRFSKELDTVDMSMPEVIMKPMGSLPNVIGACIVILLATPIAA	1116
DB	1063	IRSPMTFTPTTGVGRINRFSDDMDAVDSNLQYIFSPFFKSLTYLVTLVYLYNNMPWL	1122
QY	1117	IIIPGLIIPYVQVYVASSRQKRLSVSRSPVYSHNETLLGVSVITRAPEEORFTH	1176
DB	1123	VNFMFLVYIYQYFYVLSRELKRLISISYSPIMSLMSESLNGYSIIDADHFERFIY	1182
QY	1177	OSDLKVDENQKAYPYSIVANRWLAVRLECVGNCIVFAALFAVI---SRHSLSAGLVGUS	1233
DB	1183	LNVEKIQYNVDVFNFRSTNRWLSVRLQITIGATIVLATAILATAMTKQLSSGMVGLL	1242
QY	1234	VSYSLQVTTYLNLVRMSEMETNIVAVERLKEYSETEKEAPWQOETAAPPSWPQVGRV	1293
DB	1243	MSYSELVGTSLTWIVRTTITETNVSVERIVEYCELPPEAQSIINPEKRPDENWPSKGI	1302
QY	1294	EPNRYCLAVREDLDFVLARHINVTINGGEKVGIVGRTGAGKSLTLGLPRINSAEGEII	1353
DB	1303	EPKNYSTKYRENLDPLVANNINVKIEFCRKGIVIGRTGAGKSLTSLALFILPEKTEKII	1362
QY	1354	DGINIAKIGLHDLRFKTIIPQDPVLFSGSLRMNLDPPFSQYSDEEVTSTLELAHLKDFVS	1413

Db	1363	DGIDISDIGLFDLRSHLAIIPQDAQAFEGTVKTNLDPFNRYSEDEBKRAVEQAHLKPHLE	1422
Qy	1414	AL-----PDKLDHECAEGGENLSVGQSQLVCLARALLRKYKILVLD	1456
Db	1423	KMLHSPRGDDSNEDGNVNDILDVKINENGNSLSVGQSQLCLARALINRSKILVLD	1482
Qy	1457	TAAVDLETDDLIOSTIRTOPECTVLTIAHRLNTIMDYTRVIVLDKGEIOEYGAPSDLLQ	1516
Db	1483	TASVDMETDKIIQDTIRREFKORTILTIAHRTDVLDSDKIIVLDQGSVREDFSPSKLS	1542
Qy	1517	QR-GLFYSMAXDAG	1529
Db	1543	DKTSIFYSLCEKGG	1556

Search completed: March 18, 2005, 11:05:27  
Job time : 62.2651 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2005, 10:57:20 ; Search time 203.942 Seconds  
(without alignments)  
4838.524 Million cell updates/sec

Title: US-10-665-283-1  
Perfect score: 9903  
Sequence: 1 MALRGFCGADGSDPLDWNV.....RSVAVAKPKFSISPDLSL 1927

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt\_03:\*

1: uniprot\_sprot:\*

2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7860	79.4	1531	1 MRP1_HUMAN	P33527 homo sapien
2	7769	78.5	1515	2 Q9UQ99	Q9UQ99 homo sapien
3	7695	77.7	1531	2 Q864R9	Q864R9 macaca fasc
4	7689	77.6	1531	2 Q864S0	Q864S0 macaca fasc
5	7444	75.2	1459	2 Q9UQ97	Q9UQ97 homo sapien
6	7419.5	74.9	1456	2 Q9UQ97	Q9UQ97 homo sapien
7	7272	73.4	1531	2 Q6UR05	Q6UR05 canis famil
8	7230.5	73.0	1530	2 Q8HXQ5	Q8HXQ5 bos taurus
9	7104.5	71.7	1400	2 Q9UQ98	Q9UQ98 homo sapien
10	7002.5	70.7	1528	1 MRP1_MOUSE	O35379 mus musculu
11	6932.5	70.0	1532	2 Q810E4	Q810E4 rattus norv
12	6932.5	70.0	1532	2 Q8CG09	Q8CG09 rattus norv
13	6892	69.6	1523	2 Q810G9	Q810G9 rattus norv
14	6075	61.3	1215	2 Q86CP7	Q86CP7 homo sapien
15	4522	45.7	1519	2 Q80ZK8	Q80ZK8 mus musculu
16	4491.5	45.4	1527	1 MRP3_HUMAN	O15438 homo sapien
17	4422	44.7	1514	2 Q9G0A9	Q9G0A9 homo sapien
18	4396	44.4	1522	1 MRP3_RAT	O88563 rattus norv
19	3952.5	39.9	1592	2 Q7Q1D7	Q7Q1D7 anopheles g
20	3885.5	39.2	1548	2 Q7KTC3	Q7KTC3 drosophila
21	3880.5	39.2	1548	2 Q7KTC1	Q7KTC1 drosophila
22	3879.5	39.2	1548	2 Q7KTB7	Q7KTB7 drosophila
23	3874.5	39.1	1548	2 Q7KTC2	Q7KTC2 drosophila
24	3860.5	39.0	1548	2 Q7KTC0	Q7KTC0 drosophila
25	3857	38.9	1549	2 Q7KTC5	Q7KTC5 drosophila
26	3845.5	38.8	1548	2 Q879C5	Q879C5 drosophila
27	3838.5	38.8	1548	2 Q7KTB8	Q7KTB8 drosophila
28	3836.5	38.7	1548	2 Q7KTB9	Q7KTB9 drosophila
29	3818.5	38.6	1548	2 Q7KTD0	Q7KTD0 drosophila
30	3813.5	38.5	1548	2 Q7KTC8	Q7KTC8 drosophila
31	3812.5	38.5	1548	2 Q7KTC4	Q7KTC4 drosophila

32	3807.5	38.4	1548	2 Q7KTC9	Q7KTC9 drosophila
33	3793.5	38.3	1548	2 Q7KTC7	Q7KTC7 drosophila
34	3790	38.3	1549	2 Q9VK56	Q9VK56 drosophila
35	3771.5	38.1	1548	2 Q7KTC5	Q7KTC5 drosophila
36	3769.5	38.1	1548	2 Q7KTC6	Q7KTC6 drosophila
37	3741	37.8	1564	2 Q8QG98	Q8QG98 raja erinac
38	3738.5	37.8	1567	2 Q6PH26	Q6PH26 brachydanio
39	3655	36.9	822	2 Q9JHS0	Q9JHS0 rattus norv
40	3653.5	36.9	1544	2 Q6PSM3	Q6PSM3 canis famil
41	3641.5	36.8	1544	2 Q9SM36	Q9SM36 canis famil
42	3625.5	36.6	796	2 Q8C7V6	Q8C7V6 mus musculu
43	3595	36.3	1545	1 MRP2_HUMAN	Q92887 homo sapien
44	3594	36.3	1544	2 Q9SL75	Q9SL75 macaca mula
45	3577	36.1	1564	1 MRP2_RABIT	Q28689 o canalicul

## ALIGNMENTS

RESULT 1  
MRP1\_HUMAN  
ID\_MRP1\_HUMAN STANDARD; PRT; 1531 AA.  
AC P33527; O14819; P78419;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Multidrug resistance-associated protein 1 (ATP-binding cassette, sub-family C, member 1).  
DE family C, member 1).  
GN Name=ABCC1; Synonyms=MRP, MRP1;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=9308080; PubMed=1360704;  
RA Cole S.P.C., Bhargava J.G., Gerlach J.H., Mackie J.E., Grant C.E., Almquist K.C., Stewart A.J., Kurz E.U., Duncan A.M.V., Deeley R.G.;  
RT "Overexpression of a transporter gene in a multidrug-resistant human lung cancer cell line."  
RT Science 258:1650-1654(1992).  
RN [2]  
RP REVISIONS.  
RX MEDLINE=93262415; PubMed=8098549;  
RA Cole S.P.C., Deeley R.G.;  
RT "Multidrug resistance-associated protein: sequence correction."  
RL Science 260:879-879(1993).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96251691; PubMed=8649356;  
RA Stride B.D., Valdimarsson G., Gerlach J.H., Wilson G.M., Cole S.P.C., Deeley R.G.;  
RT "Structure and expression of the messenger RNA encoding the murine multidrug resistance protein, an ATP-binding cassette transporter."  
RL Mol. Pharmacol. 49:962-971(1996).  
RN [4]  
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.  
RX MEDLINE=98008927; PubMed=9344662; DOI=10.1006/geno.1997.4950;  
RA Grant C.E., Kurz E.U., Cole S.P.C., Deeley R.G.;  
RT "Analysis of the intron-exon organization of the human multidrug-resistance protein gene (MRP) and alternative splicing of its mRNA."  
RL Genomics 45:368-378(1997).  
RN [5]  
RP SEQUENCE OF 1131-1531 FROM N.A.  
RX MEDLINE=9942570; PubMed=10493829; DOI=10.1006/geno.1999.5927;  
RA Loftus B.J., Kim U.-J., Shedd U.P., Kalush F., Brandon R., Fuhrmann J., Mason T., Crosby M.L., Barnstead M., Cronin L., Deslattes Mays A., Cao Y., Xu R.X., Kang H.-L., Mitchell S., Eichler E.E., Harris P.C., Venter J.C., Adams M.D.;  
RT "Genome duplications and other features in 12 Mb of DNA sequence from human chromosome 16p and 16q."  
RL Genomics 60:295-308(1999).  
RN [6]

RP TOPOLOGY, AND CARBOHYDRATE-LINKAGE SITES.  
RX MEDLINE=97442425; PubMed=9295302; DOI=10.1074/jbc.272.38.23623;  
RA Hipfner D.R., Almquist K.C., Leslie E.M., Gerlach J.H., Grant C.E.,  
RA Deeley R.G., Cole S.P.C.;  
RT "Membrane topology of the multidrug resistance protein (MRP). A study  
RT of glycosylation-site mutants reveals an extracytosolic NH2  
RT terminus";  
RL J. Biol. Chem. 272:23623-23630(1997).  
RN [7]  
RP TOPOLOGY.  
RX MEDLINE=97476249; PubMed=93342225; DOI=10.1074/jbc.272.42.26479;  
RA Kast C., Gros P.;  
RT "Topology mapping of the amino-terminal half of multidrug resistance-  
RT associated protein by epitope insertion and immunofluorescence.";  
RL J. Biol. Chem. 272:26479-26487(1997).  
RN [8]  
RP TOPOLOGY.  
RX MEDLINE=98153110; PubMed=94853377; DOI=10.1021/bi972332v;  
RA Kast C., Gros P.;  
RT "Epitope insertion favors a six transmembrane domain model for the  
RT carboxy-terminal portion of the multidrug resistance-associated  
RT protein.";  
RL Biochemistry 37:2305-2313(1998).  
RN [9]  
RP MUTAGENESIS OF ASP-792; ASP-793; LYS-1333 AND 1454-ASP-GLU-1455.  
RX MEDLINE=21362977; PubMed=11469806; DOI=10.1006/abbi.2001.2441;  
RA Cui L., Hou Y.-X., Riordan J.R., Chang X.-B.;  
RT "Mutations of the Walker B motif in the first nucleotide binding  
RT domain of multidrug resistance protein MRP1 prevent conformational  
RT maturation.";  
RL Arch. Biochem. Biophys. 392:153-161(2001).  
RN [10]  
RP MUTAGENESIS OF TRP-1246  
RX MEDLINE=20283919; PubMed=11278967; DOI=10.1074/jbc.M011246200;  
RA Ito K., Olsen S.L., Qiu W., Deeley R.G., Cole S.P.C.;  
RT "Mutation of a single conserved tryptophan in multidrug resistance  
RT protein 1 (MRP1/ABCC1) results in loss of drug resistance and  
RT selective loss of organic anion transport.";  
RL J. Biol. Chem. 276:15616-15624(2001).  
RN [11]  
RP VARIANTS GLN-633 AND VAL-671.  
RX MEDLINE=20296630; PubMed=10835642; DOI=10.1038/76102;  
RA Le Saux O., Urban Z., Techetti C., Csizsar K., Bacchelli B.,  
RA Quagliari D., Pasquali-Ronchetti I., Pope F.M., Richards A., Terry S.,  
RA Bercovitch L., de Paeppe A., Boyd C.D.;  
RT "Mutations in a gene encoding an ABC transporter cause pseudoxanthoma  
RT elasticum.";  
RL Nat. Genet. 25:223-227(2000).  
RN [12]  
RP VARIANT VAL-671.  
RX MEDLINE=20283940; PubMed=10811882; DOI=10.1073/pnas.100041297;  
RA Ringpfeil F., Lebowitz M.G., Christiano A.M., Utto J.;  
RT "Pseudoxanthoma elasticum: mutations in the MRP6 gene encoding a  
RT transmembrane ATP-binding cassette (ABC) transporter.";  
RL Proc. Natl. Acad. Sci. U.S.A. 97:6001-6006(2000).  
RN [13]  
RP VARIANT SER-433, AND CHARACTERIZATION OF VARIANT VAL-671.  
RX MEDLINE=21578494; PubMed=11721885;  
RA Conrad S., Kauffmann H.-W., Ito K., Deeley R.G., Cole S.P.C.,  
RA Schrenk D.;  
RT "Identification of human multidrug resistance protein 1 (MRP1)  
RT mutations and characterization of a G671V substitution.";  
RL J. Hum. Genet. 46:656-663(2001).  
RN [14]  
RP VARIANTS MET-117 AND LEU-1512.  
RX MEDLINE=20579883; PubMed=11139250;  
RX DOI=10.1002/1098-1004(2001)17:1<74::AID-HUMU14>3.0.CO;2-F;  
RA Perdu J., Germain D.P.;  
RT "Identification of novel polymorphisms in the pms and MRP1 (ABCC1)  
RT genes at locus 16p13.1 and exclusion of both genes as responsible for  
RT pseudoxanthoma elasticum.";  
RL Hum. Mutat. 17:74-75(2001).  
RN [15]

RP VARIANTS SER-43; ILE-73; GLN-723 AND GLN-1058.  
RX MEDLINE=21163848; PubMed=11266082;  
RA Ito S., Ieiri I., Tanabe M., Suzuki A., Higuchi S., Otsubo K.;  
RT "Polymorphism of the ABC transporter genes, MDR1, MRP1 and MRP2/cMOAT,  
RT in healthy Japanese subjects.";  
RL Pharmacogenetics 11:175-184(2001).  
CC -!- FUNCTION: May participate directly in the active transport of  
CC drugs into subcellular organelles or influence drug distribution  
CC indirectly.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=8;  
CC Comment=Additional isoforms seem to exist. Experimental  
CC confirmation may be lacking for some isoforms;  
CC Name=All exons;  
CC IsoId=P33527-1; Sequence=Displayed;  
CC Name=Delexon-17;  
CC IsoId=P33527-2; Sequence=VSP\_000037;  
CC Name=Delexon-18;  
CC IsoId=P33527-3; Sequence=VSP\_000038;  
CC Name=Delexon-30;  
CC IsoId=P33527-4; Sequence=VSP\_000039;  
CC Name=Delexon-17-18;  
CC IsoId=P33527-5; Sequence=VSP\_000037, VSP\_000038;  
CC Name=Delexon-17-30;  
CC IsoId=P33527-6; Sequence=VSP\_000037, VSP\_000039;  
CC Name=Delexon-18-30;  
CC IsoId=P33527-7; Sequence=VSP\_000038, VSP\_000039;  
CC Name=Delexon-17-18-30;  
CC IsoId=P33527-8; Sequence=VSP\_000037, VSP\_000038, VSP\_000039;  
CC -!- TISSUE SPECIFICITY: Lung, testis and peripheral blood mononuclear  
CC cells.  
CC -!- SIMILARITY: Belongs to the ABC transporter family. MRP subfamily.  
CC -!- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;  
CC WWW="http://www.infobiogen.fr/services/chroncancer/Genes/MRPID106.html".  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; L05628; AAB46616.1; --  
DR EMBL; AF022853; AAB83983.1; JOINED.  
DR EMBL; AF022824; AAB83983.1; JOINED.  
DR EMBL; AF022825; AAB83983.1; JOINED.  
DR EMBL; AF022826; AAB83983.1; JOINED.  
DR EMBL; AF022827; AAB83983.1; JOINED.  
DR EMBL; AF022828; AAB83983.1; JOINED.  
DR EMBL; AF022829; AAB83983.1; JOINED.  
DR EMBL; AF022830; AAB83983.1; JOINED.  
DR EMBL; AF022831; AAB83983.1; JOINED.  
DR EMBL; AF022832; AAB83983.1; JOINED.  
DR EMBL; AF022833; AAB83983.1; JOINED.  
DR EMBL; AF022834; AAB83983.1; JOINED.  
DR EMBL; AF022835; AAB83983.1; JOINED.  
DR EMBL; AF022836; AAB83983.1; JOINED.  
DR EMBL; AF022837; AAB83983.1; JOINED.  
DR EMBL; AF022838; AAB83983.1; JOINED.  
DR EMBL; AF022839; AAB83983.1; JOINED.  
DR EMBL; AF022840; AAB83983.1; JOINED.  
DR EMBL; AF022841; AAB83983.1; JOINED.  
DR EMBL; AF022842; AAB83983.1; JOINED.  
DR EMBL; AF022843; AAB83983.1; JOINED.  
DR EMBL; AF022844; AAB83983.1; JOINED.  
DR EMBL; AF022845; AAB83983.1; JOINED.  
DR EMBL; AF022846; AAB83983.1; JOINED.  
DR EMBL; AF022847; AAB83983.1; JOINED.  
DR EMBL; AF022848; AAB83983.1; JOINED.  
DR EMBL; AF022849; AAB83983.1; JOINED.  
DR EMBL; AF022850; AAB83983.1; JOINED.

Db	961	SVTYDYMKAIGLFI	SFLSIFLFCNCHV	SALASNYWLSLWTD	DDPIVNGTQ	BEHTKVR	LSVYG	1021			
Qy	1021	ALGISQ	IAVFGYSMAV	SGIILASRCLH	VDLLHSLRSP	SPMSFERTP	SGNLVNR	FSKEL 1080			
Db	1021	ALGISQ	IAVFGYSMAV	SGIILASRCLH	VDLLHSLRSP	SPMSFERTP	SGNLVNR	FSKEL 1080			
Qy	1081	DTVDS	MIPEVIKFM	SGSLFNVCAC	IVILLATPIA	IIIPPLGLI	IVFFVORF	VVASSRQL 1140			
Db	1081	DTVDS	MIPEVIKFM	SGSLFNVCAC	IVILLATPIA	IIIPPLGLI	IVFFVORF	VVASSRQL 1140			
Qy	1141	KRL	ESVRSRSPVY	SHFNETLLG	SVIRAF	EEQERFIHQ	SDLKVDEN	KAYYPSIVANRWLA 1200			
Db	1141	KRL	ESVRSRSPVY	SHFNETLLG	SVIRAF	EEQERFIHQ	SDLKVDEN	KAYYPSIVANRWLA 1200			
Qy	1201	VR	LECVCNCIVL	PAALFAV	SRHSLAG	LVGLSVSYSLQ	VTYLNMLVR	MSSEMETNIVA 1260			
Db	1201	VR	LECVCNCIVL	PAALFAV	SRHSLAG	LVGLSVSYSLQ	VTYTYLNMLVR	MSSEMETNIVA 1260			
Qy	1261	VER	LKEYSET	KEAPWQIQ	ETAP	SPSPQVGR	VEFRNYCL	RYREDLD	DFVLRHINV	TGG 1320	
Db	1261	VER	LKEYSET	KEAPWQIQ	ETAP	SPSPQVGR	VEFRNYCL	RYREDLD	DFVLRHINV	TGG 1320	
Qy	1321	EK	GVIGVRTG	GAGKSSITL	GLFR	INESA	EGEIIID	GINIAK	IIGHDLR	FKITIIIPQDP	PVLV 1380
Db	1321	EK	GVIGVRTG	GAGKSSITL	GLFR	INESA	EGEIIID	GINIAK	IIGHDLR	FKITIIIPQDP	PVLV 1380
Qy	1381	SG	SLRNLN	DPFSQY	SDEE	VWTSLE	LHLKDF	VSALPK	DHCEA	EGENLSV	GORQLVCL 1440
Db	1381	SG	SLRNLN	DPFSQY	SDEE	VWTSLE	LHLKDF	VSALPK	DHCEA	EGENLSV	GORQLVCL 1440
Qy	1441	AR	ALLRKT	KLVLDE	ATAV	DLTDDLI	QSTIR	TQFED	CTVLIIA	HLNTIMDY	TRVIVL 1500
Db	1441	AR	ALLRKT	KLVLDE	ATAV	DLTDDLI	QSTIR	TQFED	CTVLIIA	HLNTIMDY	TRVIVL 1500
Qy	1501	DK	GEIOEY	GAPSDLL	QOQGLFY	SMAKD	AGLV	1531			
Db	1501	DK	GEIOEY	GAPSDLL	QOQGLFY	SMAKD	AGLV	1531			

RESULT 2

Q9UQ99

ID

Q9UQ99

PRELIMINARY;

PRT;

1515 AA.

AC

Q9UQ99;

DT

01-MAY-2000 (TrEMBLrel. 13, Created)

DT

01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT

01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE

Multidrug resistance protein (Fragment).

GN

Name=MRP;

OS

Homo sapiens (Human).

OC

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX

NCBI\_TaxID=9606;

RN

[1]

RP

SEQUENCE FROM N.A.

RX

MEDLINE=98008927; PubMed=9344662; DOI=10.1006/geno.1997.4950;

RA

Grant C.E., Kurz E.U., Cole S.P., Deeley R.G.;

RT

"Analysis of the intron-exon organization of the human multidrug-resistance protein gene (MRP) and alternative splicing of its mRNA.";

RL

Genomics 45:368-378(1997).

CC

-1- SIMILARITY: Belongs to the ABC transporter family.

DR

EMBL; AF022827; AAB83979.1; JOINED.

DR

EMBL; AF022828; AAB83979.1; JOINED.

DR

EMBL; AF022829; AAB83979.1; JOINED.

DR

EMBL; AF022830; AAB83979.1; JOINED.

DR

EMBL; AF022831; AAB83979.1; JOINED.

DR

EMBL; AF022832; AAB83979.1; JOINED.

DR

EMBL; AF022833; AAB83979.1; JOINED.

DR

EMBL; AF022834; AAB83979.1; JOINED.

DR

EMBL; AF022835; AAB83979.1; JOINED.

DR

EMBL; AF022836; AAB83979.1; JOINED.

DR

EMBL; AF022837; AAB83979.1; JOINED.

DR

EMBL; AF022838; AAB83979.1; JOINED.

DR

EMBL; AF022839; AAB83979.1; JOINED.

[illegible]



QY 1261 VERLKEYSETEAPQIOETAPPSPQVGRVFNRYCLRYREDLDFVLRHINTVINGG 1320  
 DB 1261 VERLKEYSETEAPQIOETAPPSPQVGRVFNRYCLRYREDLDFVLRHINTVINGG 1320  
 QY 1321 EVKGVIGRTGAGKSSLTGLFRINSAEGEIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
 DB 1321 EVKGVIGRTGAGKSSLTGLFRINSAEGEIIDGINIARIGLHDLRFKTIIPQDPVLF 1380  
 QY 1381 SGLSRNLDPFQSDSEEVWTSLELAHLKDFVSALPDKLDHECAEGGENLSVGQRLVCL 1440  
 DB 1381 SGLSRNLDPFQSDSEEVWTSLELAHLKDFVSALPDKLDHECAEGGENLSVGQRLVCL 1440  
 QY 1441 ARALLARKTKILVLDATAVLDLTDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 DB 1441 ARALLARKTKILVLDATAVLDLTDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 QY 1501 DRKEIQEYGAPSDLLQQRGLFYSMARDAGLV 1531  
 DB 1501 DRKEIQEYGAPSDLLQQRGLFYSMARDAGLV 1531

## RESULT 4

Q864S0 PRELIMINARY; PRT; 1531 AA.  
 AC Q864S0;  
 DT 01-JUN-2003 (TREMBLrel. 24, Created)  
 DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)  
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
 DE Multidrug resistance protein 1A.  
 GN Name=MRP1;  
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;  
 OC Cercopitheciae; Macaca.  
 OX NCBI\_TaxId=9541;  
 RN [1]  
 RX MEDLINE=22544876; PubMed=12657726;  
 RA Godinot N., Iversen P.W., Tabas L., Xia X., Williams D.C.,  
 RA Dantzig A.H., Perry W.L.;  
 RT "Cloning and functional characterization of the multidrug resistance-  
 associated protein (MRP1/ABCC1) from the cynomolgus monkey.";  
 RL Mol. Cancer Ther. 2:307-316(2003).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RC TISSUE=Kidney;  
 RX MEDLINE=22544876; PubMed=12657726;  
 RA Godinot N., Iversen P.W., Tabas L., Xia X., Williams D.C.,  
 RA Dantzig A.H., Perry W.L.;  
 RT "Cloning and functional characterization of the multidrug resistance-  
 associated protein (MRP1/ABCC1) from the cynomolgus monkey.";  
 RL Mol. Cancer Ther. 2:307-316(2003).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RA Perry W.L. III, Godinot N.;  
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: Belongs to the ABC transporter family.  
 DR EMBL; AY146672; AAN65348.1; -;  
 DR HSP; P08716; IMT0.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0005524; F:ATPase activity; IEA.  
 DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.  
 DR GO; GO:0000166; F:nucleotide binding; IEA.  
 DR GO; GO:0005215; F:transporter activity; IEA.  
 DR GO; GO:0006810; F:transport; IEA.  
 DR InterPro; IPR003593; AAA ATPase.  
 DR InterPro; IPR011527; ABC membrane 1.  
 DR InterPro; IPR001140; ABC TM transp.  
 DR InterPro; IPR003439; ABC transporter.  
 DR InterPro; IPR005292; MRP assoc.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR Pfam; PF00664; ABC tran; 2.  
 DR Pfam; PF00005; ABC tran; 2.  
 DR ProDom; PD000006; ABC transporter; 2.  
 DR SMART; SM00382; AAA; 2.  
 DR TIGRFAMs; TIGR00957; mrp assoc\_pro; 1.  
 DR PROSITE; PS50929; ABC TMIF; 2.  
 DR PROSITE; PS00211; ABC\_TRANSPORTER\_1; 2.  
 DR PROSITE; PS50893; ABC\_TRANSPORTER\_2; 2.

DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; UNKNOWN\_1.  
 KW ATP-binding.  
 SQ SEQUENCE 1531 AA; 171700 MW; 168712B8CC2D2B89 CRC64;  
 Query Match 77.6%; Score 7689; DB 2; Length 1531;  
 Best Local Similarity 97.6%; Pred. No. 0;  
 Matches 1494; Conservative 18; Mismatches 19; Indels 0; Gaps 0;  
 QY 1 MALRGFCADSGDPLMDWNVTWNTSNPDFTKCFQNTVLVWVPCFYLMACFPFYLYLSRH 60  
 DB 1 MALRGFCADSGDPLMDWNVTWNTSNPDFTKCFQNTVLVWVPCFYLMACFPFYLYLSRH 60  
 QY 61 DRGYIOMTPANKTKTALGFILWVCWADLFYSFWERSRGIFLAPVFLVSPFLGLITLLA 120  
 DB 61 DRGYIOMTLNKTALGFLLWVCWADLFYSFWERSRGIFLAPVFLVSPFLGLITLLA 120  
 QY 121 TFLIQLERRKGVQSSGIMLTFWLVALCALAILRSKIMTALKEDAQVDFRDIIFYVYFS 180  
 DB 121 TFLIQLERRKGVQSSGIMLTFWLVALCALAILRSKIMTALKEDAQVDFRDIIFYVYFS 180  
 QY 181 LLLIQLVLSGSDRSPFSETHDPNCPSESSASFLSRITFWWITGLIVRGVQPLEGSD 240  
 DB 181 LVLIQLVLSGSDRSPFSETHDPNCPSESSASFLSRITFWWITGLIVRGVQPLEGSD 240  
 QY 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKVYSSKDPAPQKSSKYDANEVEAL 300  
 DB 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKVYSSKDPAPQKSSKYDANEVEAL 300  
 QY 301 IVKSPQKEWNPSPFKVLYKTFPGYFLMSFFFKAIHDLMMFSGPQILKLIKFNVDTKAPD 360  
 DB 301 IVKSPQKEWNPSPFKVLYKTFPGYFLMSFFFKAIHDLMMFSGPQILKLIKFNVDTKAPD 360  
 QY 361 WQGYFYVLLFVTVACLOTLVHLQYFHICFVSGMRKTAIVGAVYRKALVITNSARKSTV 420  
 DB 361 WQGYFYTALLFVAACLOTLVHLQYFHICFVSGMRKTAIVGAVYRKALVITNSARKSTV 420  
 QY 421 GEIVNLSVDAQRFMDLATYINMIWSAPLOVILALYLLMLNLGSPVLGAVMVLMPVN 480  
 DB 421 GEIVNLSVDAQRFMDLATYINMIWSAPLOVILALYLLMLNLGSPVLGAVMVLMPVN 480  
 QY 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLYAWELAFKDKVLAIROBELKVLK 540  
 DB 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLYAWELAFKDKVLAIROBELKVLK 540  
 QY 541 KSAYLSAVGFTWVCTPFLVALCTFAVYVITDENNILDAAQTAFAVSLALFNLRPLNLP 600  
 DB 541 KSAYLSAVGFTWVCTPFLVALCTFAVYVITDENNILDAAQTAFAVSLALFNLRPLNLP 600  
 QY 601 MVISSIVQASVSLKRLRIFLSHEELEPDSITERRPVKGGTNSITVRNATFWARSDPPT 660  
 DB 601 MVISSIVQASVSLKRLRIFLSHEELEPDSITERRPVKGGTNSITVRNATFWARSDPPT 660  
 QY 661 LNGITFSIPEGALVAVVGVGCGKSSLLSALLAEMDKVEGHVAIKGSVAVYVQQAQI 720  
 DB 661 LNGITFSIPEGALVAVVGVGCGKSSLLSALLAEMDKVEGHVAIKGSVAVYVQQAQI 720  
 QY 721 SLRNLFGCOLEPYPYRSVIQACALLPDLEILPSGDRTEIGKGNLSGCKORVSLAR 780  
 DB 721 SLQENILFGCOLEPYPYRSVIQACALLPDLEILPSGDRTEIGKGNLSGCKORVSLAR 780  
 QY 781 AVSNADIYLFDDPLSAVDAHVGKHIENVIKPGMLKNKTRILVTHSMSYLPQVDVLI 840  
 DB 781 AVYCNADIYLFDDPLSAVDAHVGKHIENVIKPGMLKNKTRILVTHSMSYLPQVDVLI 840  
 QY 841 MSGGKISEMGSYQELLARDGAFELRTYASTEQDAENGVTGVSQPGKEAKQEMNGM 900  
 DB 841 MSGGKISEMGSYQELLARDGAFELRTYASTEQDAENGVTGVSQPGKEAKQEMNGM 900  
 QY 901 LVTDGAKQLOQLSSSSSYSGDISRHHNSTAELOKAEKEETWKLMEADKATQGVKL 960  
 DB 901 LVTDGAKQLOQLSSSSSYSGDVSRHHNSTAELOKAEKEETWKLMEADKATQGVKL 960  
 QY 961 SVYWDYMKAIKGLFISFLSIIFLMCNHVSALASVWLSLWTDTPVNGTQHTKVLRSVYG 1020

```

Db      961 SVYDYMKAIGLFISPLSIFLFCNHWAAASNYLWLTDDPVGNGTOHTKVLRSVYG 1020
QY      1021 ALGISGIAVGVSMVAISGGILASRCLHVDLLHSILRSPMFFERTPSGNLVNRFKSL 1080
Db      1021 ALGISGIAVGVSMVAISGGILASRCLHVDLLHSILRSPMFFERTPSGNLVNRFKSL 1080
QY      1081 DTVDMSIPEVIMKMGSLFNVCACIVILLATPIAAIIPIPLGLIYFFVQRFVYASSRQL 1140
Db      1081 DTVDMSIPEVIMKMGSLFNVCACIVILLATPIAAIIPIPLGLIYFFVQRFVYASSRQL 1140
QY      1141 KLESVSRSPVYSHFNETHLLGVSIVRAPEQERFIHQSDLKVDENQKAYPYSIVANRWLA 1200
Db      1141 KLESVSRSPVYSHFNETHLLGVSIVRAPEQERFIHQSDLKVDENQKAYPYSIVANRWLA 1200
QY      1201 VRLCEVGNICIVLFAALFAVISRHSLSAGLVLSVSYSLQVTTYLNLVMSSEMETNIVA 1260
Db      1201 VRLCEVGNICIVLFAALFAVISRHSLSAGLVLSVSYSLQVTTYLNLVMSSEMETNIVA 1260
QY      1261 VERLKEYSETEAPMQIQTAPPSPQVGRVFNRYCLRYREDLDFVLRHINVTINGG 1320
Db      1261 VERLKEYSETEAPMQIQTAPPSPQVGRVFNRYCLRYREDLDFVLRHINVTINGG 1320
QY      1321 EKVGIVGRTGAGKSSLTGLFRINESAGEIIIDGINIAKIGHDLRFKTIIPQDPVLF 1380
Db      1321 EKVGIVGRTGAGKSSLTGLFRINESAGEIIIDGINIAKIGHDLRFKTIIPQDPVLF 1380
QY      1381 SGLSRNLNLPFFSOYSDDEEVTSLAHLKDFVSALPDKLDHCAEGENLSVQRLVCL 1440
Db      1381 SGLSRNLNLPFFSOYSDDEEVTSLAHLKDFVSALPDKLDHCAEGENLSVQRLVCL 1440
QY      1441 ARALLRKTKILVDEATAVLETDLIQSTIRTOPEDCTVLTIAHRLNTIMDYTRVIVL 1500
Db      1441 ARALLRKTKILVDEATAVLETDLIQSTIRTOPEDCTVLTIAHRLNTIMDYTRVIVL 1500
QY      1501 DRGETQYEGAPSDLLQOQGLFYSMAKDAGLV 1531
Db      1501 DRGETQYEGAPSDLLQOQGLFYNMARDAGLV 1531

RESULT 5
Q9UQ97 PRELIMINARY; PRT: 1459 AA.
AC Q9UQ97;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2004 (TrEMBLrel. 26, Last annotation update)
DE Multidrug resistance protein (fragment).
GN Name=MRP;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98008927; PubMed=9344662; DOI=10.1006/geno.1997.4950;
RA Grant C.E., Kurz E.U., Cole S.P., Deeley R.G.;
RT "Analysis of the intron-exon organization of the human multidrug-
RT resistance protein gene (MRP) and alternative splicing of its mRNA.";
RL Genomics 45:368-378(1997).
CC -!- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AF022824; AAB83980.1; JOINED.
DR EMBL; AF022825; AAB83980.1; JOINED.
DR EMBL; AF022826; AAB83980.1; JOINED.
DR EMBL; AF022827; AAB83980.1; JOINED.
DR EMBL; AF022828; AAB83980.1; JOINED.
DR EMBL; AF022829; AAB83980.1; JOINED.
DR EMBL; AF022830; AAB83980.1; JOINED.
DR EMBL; AF022831; AAB83980.1; JOINED.
DR EMBL; AF022832; AAB83980.1; JOINED.
DR EMBL; AF022833; AAB83980.1; JOINED.
DR EMBL; AF022834; AAB83980.1; JOINED.
DR EMBL; AF022835; AAB83980.1; JOINED.

```

```

DR EMBL; AF022836; AAB83980.1; JOINED.
DR EMBL; AF022837; AAB83980.1; JOINED.
DR EMBL; AF022838; AAB83980.1; JOINED.
DR EMBL; AF022839; AAB83980.1; JOINED.
DR EMBL; AF022841; AAB83980.1; JOINED.
DR EMBL; AF022842; AAB83980.1; JOINED.
DR EMBL; AF022843; AAB83980.1; JOINED.
DR EMBL; AF022844; AAB83980.1; JOINED.
DR EMBL; AF022845; AAB83980.1; JOINED.
DR EMBL; AF022846; AAB83980.1; JOINED.
DR EMBL; AF022847; AAB83980.1; JOINED.
DR EMBL; AF022848; AAB83980.1; JOINED.
DR EMBL; AF022849; AAB83980.1; JOINED.
DR EMBL; AF022850; AAB83980.1; JOINED.
DR EMBL; AF022851; AAB83980.1; JOINED.
DR EMBL; AF022852; AAB83980.1; JOINED.
DR EMBL; AF022853; AAB83980.1; JOINED.
DR HSSP; P08716; 1MT0.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0005524; F: ATP binding; IEA.
DR GO; GO:0042626; F: ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO; GO:0000166; F: nucleotide binding; IEA.
DR GO; GO:0005215; F: transporter activity; IEA.
DR GO; GO:0006810; F: transport; IEA.
DR InterPro; IPR003593; AAA ATPase.
DR InterPro; IPR011527; ABC membrane 1.
DR InterPro; IPR011140; ABC TM transp.
DR InterPro; IPR003439; ABC transporter.
DR InterPro; IPR005292; MRP_assoc.
DR Pfam; PF00664; ABC_tran; 1.
DR Pfam; PF00664; ABC_tran; 2.
DR ProDom; PD000006; ABC_transpouter; 1.
DR SMART; SM00382; AAA; 2.
DR TIGRfam; TIGR00957; MRP_assoc_pro; 1.
DR PROSITE; PS00929; ABC_TMIF; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; UNKNOWN_1.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
KW ATP-binding.
FT NON_TER 1
SQ SEQUENCE 1459 AA; 163830 MW; 6A016ASA6E7AA95C CRC64;

Query Match 75.2%; Score 7444; DB 2; Length 1459;
Best Local Similarity 96.3%; Pred. No. 0;
Matches 1459; Conservative 0; Mismatches 0; Indels 56; Gaps 1;

QY 17 DWNVTWNTSNPDFTKCFONTVLVWVPCFYLWACFPFFLYLSRHRDGYIQMTPLNKTITA 76
Db 1 DWNVTWNTSNPDFTKCFONTVLVWVPCFYLWACFPFFLYLSRHRDGYIQMTPLNKTITA 60
QY 77 LGFLLMIVCWADLFYSFWERSRGIFLAPVFLVSPDLLGTTLLATFLQLERRKGVQSSG 136
Db 61 LGFLLMIVCWADLFYSFWERSRGIFLAPVFLVSPDLLGTTLLATFLQLERRKGVQSSG 120
QY 137 IMLTFLWALVCAILALSKIMTALKEDAQVDLFRDITFYVYFSLLLQLVLSCFSDRSP 196
Db 121 IMLTFLWALVCAILALSKIMTALKEDAQVDLFRDITFYVYFSLLLQLVLSCFSDRSP 180
QY 197 LFSETHDPNCPSSASFLSRITFWITGLIVRGYRPLEGSDLSLWLNKETSQVVPV 256
Db 181 LFSETHDPNCPSSASFLSRITFWITGLIVRGYRPLEGSDLSLWLNKETSQVVPV 240
QY 257 LVKNWKKCAKTRKQPVKVYVSSKPAQPKSSKVDANEVEEALIVKSPQKWNPSLFKV 316
Db 241 LVKNWKKCAKTRKQPVKVYVSSKPAQPKSSKVDANEVEEALIVKSPQKWNPSLFKV 300
QY 317 LYKTGPGYFLMSFFFKAIHDLMMFSGPQLIKLIFVNDTKAPDQWGYFYTVLLFVTACL 376
Db 301 LYKTGPGYFLMSFFFKAIHDLMMFSGPQLIKLIFVNDTKAPDQWGYFYTVLLFVTACL 360
QY 377 QTLVLHGYFHICFVSGMKIKTAVIGAVYRKALVIYNSARKSVTGEIVNLMSVDQAPFMD 436

```

Db 361 QTLVHLQYFHCIVSGMIRIKTAVIGAYRKALVITNSARKSVTGEIVNLSVDAQRFMD 420  
 QY 437 LATYINMIWSAPLOVTLALYLWNLGSPSVLAGVAVMLVMPVNAVAMKTKTYQVAHMK 496  
 Db 421 LATYINMIWSAPLOVTLALYLWNLGSPSVLAGVAVMLVMPVNAVAMKTKTYQVAHMK 480  
 QY 497 SKDNRIKLMNEILNGIKVLKXAWELAFKDKVLAIRQBELKVLKKSAYLSAVGTFTWCT 556  
 Db 481 SKDNRIKLMNEILNGIKVLKXAWELAFKDKVLAIRQBELKVLKKSAYLSAVGTFTWCT 540  
 QY 557 PELVALCTFAVVTIDENNILDAQTAFAVSLALFNILRPINILPMVSIQASVSLKRL 616  
 Db 541 PELVALCTFAVVTIDENNILDAQTAFAVSLALFNILRPINILPMVSIQASVSLKRL 600  
 QY 617 RIFLSEHELEPPSIERRPVKDGSGTNSITVRNATFTWASDPTGLNGITFSPGALYAV 676  
 Db 601 RIFLSEHELEPPSIERRPVKDGSGTNSITVRNATFTWASDPTGLNGITFSPGALYAV 660  
 QY 677 VQVCGKSSLSALLAEMDKVEGHVAIKGSVAYVPQQAQWIONDSLRENILFGCOLEBPY 736  
 Db 661 VQVCGKSSLSALLAEMDKVEGHVAIKGSVAYVPQQAQWIONDSLRENILFGCOLEBPY 720  
 QY 737 YRSVTOACALLPDLEILPSGDRTEIGEKGVNLSGGQKQKRVSLARAVYSNADIIYLFDDPLS 796  
 Db 721 YRSVTOACALLPDLEILPSGDRTEIGEKGVNLSGGQKQKRVSLARAVYSNADIIYLFDDPLS 747  
 QY 797 AVDAHVGKHI FENVIGPKGMLKNKTRILVTHSMSYLPQVDVLIIVMSGGKISMGYSQELL 856  
 Db 748 AVDAHVGKHI FENVIGPKGMLKNKTRILVTHSMSYLPQVDVLIIVMSGGKISMGYSQELL 784  
 QY 857 ARDGAFAEFLRTYASTEQEQDAEENGVTGSGPGKEAKOMENGLVTPDSAGKQLQRLSS 916  
 Db 785 ARDGAFAEFLRTYASTEQEQDAEENGVTGSGPGKEAKOMENGLVTPDSAGKQLQRLSS 844  
 QY 917 SSSYSGDISRHNSAEKQAKKEETWKLMEADKAQGVKLSVYWDYKKAIGLFTSF 976  
 Db 845 SSSYSGDISRHNSAEKQAKKEETWKLMEADKAQGVKLSVYWDYKKAIGLFTSF 904  
 QY 977 LSIFLFMCNHSALASNTYLSLWTDPIVNGTQHTKVLRSVYGALGISQGIYAVFGYSMA 1036  
 Db 905 LSIFLFMCNHSALASNTYLSLWTDPIVNGTQHTKVLRSVYGALGISQGIYAVFGYSMA 964  
 QY 1037 VSIGGILASRCLHVDLLSHILRSPMSFERTPTSGNLVNRFSKELTVDSDMPEVIKMPMG 1096  
 Db 965 VSIGGILASRCLHVDLLSHILRSPMSFERTPTSGNLVNRFSKELTVDSDMPEVIKMPMG 1024  
 QY 1097 SLFNIVGACIVLLATPTAAIIIPGLIYFPVQRFYVASSRQLKLESVSRSPYSHFN 1156  
 Db 1025 SLFNIVGACIVLLATPTAAIIIPGLIYFPVQRFYVASSRQLKLESVSRSPYSHFN 1084  
 QY 1157 ETLGLSVIRAFEEQERFIHQSDLKVDENQKAYPSIVANRWLAVRLECVGNCIVLFAAL 1216  
 Db 1085 ETLGLSVIRAFEEQERFIHQSDLKVDENQKAYPSIVANRWLAVRLECVGNCIVLFAAL 1144  
 QY 1217 FAVISRHSLSAGLVGLSVYSQVTTYNLVLRMSSEMETNIVAVERLKEYSETEKEAPW 1276  
 Db 1145 FAVISRHSLSAGLVGLSVYSQVTTYNLVLRMSSEMETNIVAVERLKEYSETEKEAPW 1204  
 QY 1277 QIQTAPSPSPQVGRVFRNYCLRYREDLPVLRHNTINGEKVIGVORTGAGKSSL 1336  
 Db 1205 QIQTAPSPSPQVGRVFRNYCLRYREDLPVLRHNTINGEKVIGVORTGAGKSSL 1264  
 QY 1337 TIGLFRINESARGEIIIDGINIAKIGLHDLRFKTIIPQDPVLFSGSLRMNLDPFPSQYSD 1396  
 Db 1265 TIGLFRINESARGEIIIDGINIAKIGLHDLRFKTIIPQDPVLFSGSLRMNLDPFPSQYSD 1324  
 QY 1397 EEVWTSLELAHLKDFVSLPDKLDHECAEGGENLSVGQRLVCLARALLRKTILVLDEA 1456  
 Db 1325 EEVWTSLELAHLKDFVSLPDKLDHECAEGGENLSVGQRLVCLARALLRKTILVLDEA 1384  
 QY 1457 TAAVLETDLLIQSTIRTFQFEDCTVLTIAHRLNTIMDYTRIVLDKGEIQYVAGPSDLLQ 1516  
 Db 1385 TAAVLETDLLIQSTIRTFQFEDCTVLTIAHRLNTIMDYTRIVLDKGEIQYVAGPSDLLQ 1444

QY 1517 ORGLFYSNAKDAGLV 1531  
 Db 1445 ORGLFYSNAKDAGLV 1459  
 RESULT 6  
 Q9UQA0  
 ID Q9UQA0 PRELIMINARY; PRT; 1456 AA.  
 AC Q9UQA0;  
 DT 01-MAY-2000 (Tremblrel. 13, Created)  
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
 DT 01-MAY-2004 (Tremblrel. 26, Last annotation update)  
 DE Multidrug resistance protein (Fragment).  
 GN Name=MRP;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98008927; PubMed=9344662; DOI=10.1006/geno.1997.4950;  
 RA Grant C.E., Kurz E.U., Cole S.P., Deeley R.G.;  
 RT "Analysis of the intron-exon organization of the human multidrug-  
 resistance protein gene (MRP) and alternative splicing of its mRNA.";  
 RL Genomics 45:368-378(1997).  
 CC -1- SIMILARITY: Belongs to the ABC transporter family.  
 DR EMBL; AF022824; AAB83981.1; JOINED.  
 DR EMBL; AF022825; AAB83981.1; JOINED.  
 DR EMBL; AF022826; AAB83981.1; JOINED.  
 DR EMBL; AF022827; AAB83981.1; JOINED.  
 DR EMBL; AF022828; AAB83981.1; JOINED.  
 DR EMBL; AF022829; AAB83981.1; JOINED.  
 DR EMBL; AF022830; AAB83981.1; JOINED.  
 DR EMBL; AF022831; AAB83981.1; JOINED.  
 DR EMBL; AF022832; AAB83981.1; JOINED.  
 DR EMBL; AF022833; AAB83981.1; JOINED.  
 DR EMBL; AF022834; AAB83981.1; JOINED.  
 DR EMBL; AF022835; AAB83981.1; JOINED.  
 DR EMBL; AF022836; AAB83981.1; JOINED.  
 DR EMBL; AF022837; AAB83981.1; JOINED.  
 DR EMBL; AF022838; AAB83981.1; JOINED.  
 DR EMBL; AF022840; AAB83981.1; JOINED.  
 DR EMBL; AF022841; AAB83981.1; JOINED.  
 DR EMBL; AF022842; AAB83981.1; JOINED.  
 DR EMBL; AF022843; AAB83981.1; JOINED.  
 DR EMBL; AF022844; AAB83981.1; JOINED.  
 DR EMBL; AF022845; AAB83981.1; JOINED.  
 DR EMBL; AF022846; AAB83981.1; JOINED.  
 DR EMBL; AF022847; AAB83981.1; JOINED.  
 DR EMBL; AF022848; AAB83981.1; JOINED.  
 DR EMBL; AF022849; AAB83981.1; JOINED.  
 DR EMBL; AF022850; AAB83981.1; JOINED.  
 DR EMBL; AF022851; AAB83981.1; JOINED.  
 DR EMBL; AF022852; AAB83981.1; JOINED.  
 DR EMBL; AF022853; AAB83981.1; JOINED.  
 DR HSP; P08716; 1WTO.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m...; IEA.  
 DR GO; GO:0000166; F:nucleotide binding; IEA.  
 DR GO; GO:0005215; F:transporter activity; IEA.  
 DR GO; GO:0006810; P:transport; IEA.  
 DR InterPro; IPR003593; AAA\_ATPase.  
 DR InterPro; IPR011527; ABC\_membrane\_1.  
 DR InterPro; IPR001140; ABC\_TM\_transp.  
 DR InterPro; IPR003439; ABC\_transporter.  
 DR InterPro; IPR005292; MRP\_assoc.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR Pfam; PF00664; ABC\_membrane; 2.  
 DR Pfam; PF00005; ABC\_tran; 2.  
 DR ProDom; PD000006; ABC\_transporter; 2.  
 DR SMART; SM00382; AAA; 2.

```
DR TIGR00957; MRP_assoc_pro; 1.
DR PROSITE; P500929; ABC_TMIF; 2.
DR PROSITE; P500211; ABC_TRANSPORTER_1; 2.
DR PROSITE; P500893; ABC_TRANSPORTER_2; 2.
DR PROSITE; P500107; PROTEIN_KINASE_ATP; UNKNOWN_1.
KW ATP-binding.
FT NON_TER
SQ SEQUENCE 1456 AA; 163231 MW; 8DE8AAB22BC481P2 CRC64;
Query Match 74.9%; Score 7419.5; DB 2; Length 1456;
Best Local Similarity 96.1%; Pred. No. 0;
Matches 1456; Conservative 0; Mismatches 0; Indels 59; Gaps 1;
QY 17 DNNVTWNTSNPDFTKCFQNTLVWVPCFYLMACFFPYFLYLSHRDRGYQMTPLNKTKA 76
Db 1 DNNVTWNTSNPDFTKCFQNTLVWVPCFYLMACFFPYFLYLSHRDRGYQMTPLNKTKA 60
QY 77 LGFLWIVCWADLFYSFWSRSGIFLAPVFLVSPILLGITTLATFLQLERRKGVQSSG 136
Db 61 LGFLWIVCWADLFYSFWSRSGIFLAPVFLVSPILLGITTLATFLQLERRKGVQSSG 120
QY 137 IMLTFWLVALVACALALRSKIMTALKEDAQVDFRDIITYVYFSLLLQLVLSCFSDRSP 196
Db 121 IMLTFWLVALVACALALRSKIMTALKEDAQVDFRDIITYVYFSLLLQLVLSCFSDRSP 180
QY 197 LFSETHDNPCESSASFLSRITFWITGLIVRGYRQPLEGSDLSLNKEDTSEQVVPV 256
Db 181 LFSETHDNPCESSASFLSRITFWITGLIVRGYRQPLEGSDLSLNKEDTSEQVVPV 240
QY 257 LVNKKKECAKTRKOPKPVVYSSKDPAPQKESKVDANEVEEALIVKSPQKEWNPFLFKV 316
Db 241 LVNKKKECAKTRKOPKPVVYSSKDPAPQKESKVDANEVEEALIVKSPQKEWNPFLFKV 300
QY 317 LYKTFGPYELMGFFFAHDLMMFSGPOLKLLIKFVNDTKAPDQGVYTYVLLFTACL 376
Db 301 LYKTFGPYELMGFFFAHDLMMFSGPOLKLLIKFVNDTKAPDQGVYTYVLLFTACL 360
QY 377 QTLVLHQYFHCIFVSGMRITKAVIGAVYRKALVITNSARKSSTVGEIVNLMVSDAQRMD 436
Db 361 QTLVLHQYFHCIFVSGMRITKAVIGAVYRKALVITNSARKSSTVGEIVNLMVSDAQRMD 420
QY 437 LATYINMIWSAPLOVITLALYLWNLGSPVLGAVVMVLPVNAVMAKTKTYVAHMK 496
Db 421 LATYINMIWSAPLOVITLALYLWNLGSPVLGAVVMVLPVNAVMAKTKTYVAHMK 480
QY 497 SKDNRIKLMNEILNGIKVLKYAWELAFKDKVLAIRQBELKYLKKSAYLSAVGTFTWCT 556
Db 481 SKDNRIKLMNEILNGIKVLKYAWELAFKDKVLAIRQBELKYLKKSAYLSAVGTFTWCT 540
QY 557 PELVALCTPAVVTIDENNILDAQTAFAVSLALFNILRPPLNIPWVISSIVQASVSLKRL 616
Db 541 PELVALCTPAVVTIDENNILDAQTAFAVSLALFNILRPPLNIPWVISSIVQASVSLKRL 600
QY 617 RIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWASDPPTLNGITFTSPEGALVAV 676
Db 601 RIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWASDPPTLNGITFTSPEGALVAV 660
QY 677 VQGVCGGKSSILLSALLAENDKVEGHVAIKGSVAYYVQQAWIQNDSIRENILEGCOLEBPY 736
Db 661 VQGVCGGKSSILLSALLAENDKVEGHVAI 688
QY 737 YRSVIQACALLPDLIELPSGDRTEIGEKGWNLGGOKORVSLARAVYSNADYLFDDPLS 796
Db 689 -----KGVNLGGOKORVSLARAVYSNADYLFDDPLS 721
QY 797 AVDAHVGKHFENVIPGKMLKNKTRILVTHSMYLPQVDVITVMSGGKISEMGVQELL 856
Db 722 AVDAHVGKHFENVIPGKMLKNKTRILVTHSMYLPQVDVITVMSGGKISEMGVQELL 781
QY 857 ARDGAFAELRYASTEQDAEENGVTGSGPGKEAKOMENGLVTDASAGLORQLSS 916
Db 782 ARDGAFAELRYASTEQDAEENGVTGSGPGKEAKOMENGLVTDASAGLORQLSS 841
```

## RESULT 7

```
QUR05 ID Q6UR05 PRELIMINARY; PRT; 1531 AA.
AC Q6UR05;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Multidrug resistance-associated protein 1.
GN Name=MRP1;
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22404678; PubMed=13516967;
RA Ma L., Pratt S.E., Cao J., Dantzig A.H., Moore R.E., Slapak C.A.;
RT "Identification and characterization of the canine multidrug
RT resistance-associated protein."
RL Mol. Cancer Ther. 1:1335-1342(2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Ma L., Pratt S.E., Cao J., Dantzig A.H., Moore R.E., Slapak C.A.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AY363728; AAQ23148.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
```

DR GO: 0042626; F:ATPase activity, coupled to transmembrane m. . . ; IEA.  
 DR GO: 0000166; F:nucleotide binding; IEA.  
 DR GO: 0005215; F:transporter activity; IEA.  
 DR GO: 0008610; F:transport; IEA.  
 DR InterPro: IPR003593; AAA ATPase.  
 DR InterPro: IPR011527; ABC membrane 1.  
 DR InterPro: IPR001140; ABC TM transp.  
 DR InterPro: IPR003439; ABC transporter.  
 DR InterPro: IPR001395; ABC/ket\_red.  
 DR InterPro: IPR005292; MRP\_assoc.  
 DR InterPro: IPR000719; Prot\_kinase.  
 DR Pfam: PF00664; ABC membrane; 2.  
 DR Pfam: PF00005; ABC trans; 2.  
 DR ProDom: PD000006; ABC transporter; 2.  
 DR SMART: SM00382; AAA; 2.  
 DR TIGRfam: TIGR00957; MRP\_assoc\_pro; 1.  
 DR PROSITE: PS00929; ABC TMIF; 2.  
 DR PROSITE: PS00211; ABC TRANSPORTER\_1; 2.  
 DR PROSITE: PS00893; ABC TRANSPORTER\_2; 2.  
 DR PROSITE: PS00063; ALDO-KETO REDUCTASE 3; UNKNOWN\_1.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; UNKNOWN\_1.  
 DR ATP-binding.  
 DR SEQUENCE 1531 AA; 171790 MW; 746361A71C6158BD CRC64;  
 SQ

Query Match 73.4%; Score 7272; DB 2; Length 1531;  
 Best Local Similarity 91.6%; Pred. No. 0;  
 Matches 1402; Conservative 70; Mismatches 59; Indels 0; Gaps 0;

QY 1 MALRFGCSADGSDPLWNNVTWNTNPDFTKCFQNTLVWVPCFYLWACFPFYFLYLSRH 60  
 DB 1 MALRFGCSADGSDPFWEDVSNWNTNPDFTKCFQNTLVWVPCFYLWACFPFYFLYLSRH 60

QY 61 DRYGTYQMTPLNKTALGFLWIVCWADLFLYFWERSGIFLAPVFLYPSLTLLGTTLLA 120  
 DB 61 DRYGTYQMTPLNKTALGFLWIVCWADLFLYFWERSGIFLAPVFLYPSLTLLGTTLLA 120

QY 121 TFLIQLERRKGVSQSGIMLTFWLVALVICALALIRSKIMTALKEDAQVDLFRDITFYVYFS 180  
 DB 121 TFLIQLERRKGVSQSGIMLTFWLVALVICALALIRSKIMTALKEDAEIDVFRDITFYVYFS 180

QY 181 LLLIQLVLSFSDRPLSETHDNPCESSASFLSRITFWITGLIVRGVROPLEGSD 240  
 DB 181 LLLIQLVLSFSDRPLSETHDNPCESSASFLSRITFWITGLIVRGVROPLEGSD 240

QY 241 LWSLNKEDTSEQVPLVKNWKECAKTRKQPVVYSSKOPAPKESKVDANEEVEAL 300  
 DB 241 LWSLNKEDTSEQVPLVKNWKECAKTRKQPVVYSSKOPAPKESKVDANEEVEAL 300

QY 301 IVKSPQKWNPSLFKVLKTYTGPYFLMSFFFKAIHDLMMFSGPQLIKLLIKFVNDTKAPD 360  
 DB 301 IVKTPKEREPSLFKVLKTYTGPYFLMSFFFKALHDLMMFAGPEILKLLINFVNDKKAPD 360

QY 361 WQGYFYTVLLFTACLOTILVHOYPHICFVSGMRITKAVIGAVYKALVITNSARKSTV 420  
 DB 361 WQGYLYTALLFTACLOTILVHOYPHICFVSGMRITKAVIGAVYKALVITNSARKSTV 420

QY 421 GEIVNLMSVDAORFMDLTYINMWSAPLOVTLALYLWNLGSPVLGAVVWMLMPVN 480  
 DB 421 GEIVNLMSVDAORFMDLTYINMWSAPLOVTLALYLWNLGSPVLGAVVWMLMPVN 480

QY 481 AVAMAKTKTYQVAHMKSKDNRIKLMEILNGIKVLKYAWELAFKDKVLAIROBELKVLK 540  
 DB 481 AVAMAKTKTYQVAHMKSKDNRIKLMEILNGIKVLKYAWELAFKDKVLAIROBELKVLK 540

QY 541 KSAYLSAVGTFTWCTPPLVALCTFAVVTVDENNILDAQTAFAVSLALFNILRPLNLP 600  
 DB 541 KSAYLSAVGTFTWCTPPLVALCTFAVVTVDENNILDAQTAFAVSLALFNILRPLNLP 600

QY 601 MVISSIVQASVSLKRLRIFLGSHEELEDPSIERRPVKDGGGNTSITVRNATFTTWARDSDPT 660  
 DB 601 MVISSIVQASVSLKRLRIFLGSHEELEDPSIERRPVKDGGGNTSITVRNATFTTWARDSDPT 660

QY 661 LNTGITSPEGALVAVGVQCGKSSLLSALLAEMDKVEGHVAIKGSVAYVPPQQAWIQND 720

DB 661 LNTGITSPEGALVAVGVQCGKSSLLSALLAEMDKVEGHVAIKGSVAYVPPQQAWIQND 720  
 QY 721 SLRENILFGCOLEBEPYRSVIOACALLPDLBILPSGDRTEIGEKGWNLGGQKQKRVSLAR 780  
 DB 721 SLRENILFGCOLEBEPYRSVIOACALLPDLBILPSGDRTEIGEKGWNLGGQKQKRVSLAR 780  
 QY 781 AVYSNADIYLFDDPLSAVDHVGKHFENVIKPGKMLKNKTRILVTHSMVSLPOVDVILV 840  
 DB 781 AVYCDSDIYLFDDPLSAVDHVGKHFENVIKPGKMLKNKTRILVTHSMVSLPOVDVILV 840

QY 841 MSGGKISEMGYSQELLARDGAPAFELRYASTEOEDAEENGVTGVSFGPKGKAKOMENG 900  
 DB 841 MTGGKISEMGYSQELLARDGAPAFELRYASGDQEAQDDGLTGVSFGPKGKAKOMENG 900

QY 901 LVTDAGKQLOQLSSSSSYSGDISRHHNSTAELOKABKKEETWKLMEADKAQGTQVKL 960  
 DB 901 LVTDVAGKQLOQLSSSSSYSGDVSRRHTSTAELOKABKKEETWKLMEADKAQGTQVKL 960

QY 961 SVYDYMKAIGLFIISFLPMCNHVSALASNYWLSLWTDPIVNGTQEHKTVLSVYG 1020  
 DB 961 SVYDYMKAIGLFIISFLPMCNHVSALASNYWLSLWTDPIVNGTQEHKTVLSVYG 1020

QY 1021 ALGISQGIADVFGYSMAVSIIGGILASRCILHVDLLHSILRSPMSFFERTPSGNLVNRFPSKEL 1080  
 DB 1021 ALGISQGITVFGYSMAVSIIGGIFASRRLLHVDLLQNLVLRSPMSFFERTPSGNLVNRFPSKEL 1080

QY 1081 DTVDMSIPEVTKPMGSLFNIVGACIVILLATPIAAIIIPPLGLIYFFVQFYVASSRQL 1140  
 DB 1081 DTVDMSIPQVTKPMGSLFNIVGACIIILLATPIASIIIPPLGLIYFFVQFYVASSRQL 1140

QY 1141 KRLESVRSVPYSHFNETLLGVSVIRAFEEQERFIQSDLKVDENQKAYPSIVANRWLA 1200  
 DB 1141 KRLESVRSVPYSHFNETLLGVSVIRAFEEQERFIQSDLKVDENQKAYPSIVANRWLA 1200

QY 1201 VRLECVGNCIVLFAALFAVISRHSLSAGLVGSYSLSQVTTYLNWLVMSSEMETNIVA 1260  
 DB 1201 VRLECVGNCIVLFAALFSVISRHSLSAGLVGSYSLSQVTTYLNWLVMSSEMETNIVA 1260

QY 1261 VERLKEYSETKEAPWQIQETAPPSSWPQVGRVFRNYCLARYEDDLVLRHINVTINGG 1320  
 DB 1261 VERLKEYSETKEAPWQIQEMAPPSTWQVGRVFRDYGLRYRENLDLVLKHINVTINGG 1320

QY 1321 EKVGIKGTGAGKSLTGLFRINESAGEIIIDGINAKIGLHDLRKTITIPDDPVLF 1380  
 DB 1321 EKVGIKGTGAGKSLTGLFRINESAGEIIIDGINAKIGLHDLRKTITIPDDPVLF 1380

QY 1381 SGSLRMNLDPFSQYSDEEVMTSLELAHLKDFVSALPKDLHDECAEGGNSLVGQRLVCL 1440  
 DB 1381 SGSLRMNLDPFSQYSDEEVMTSLELAHLKDFVSALPKDLHDECAEGGNSLVGQRLVCL 1440

QY 1441 ARALLRTKILVLDEATAAVIDETDDLIQSTIRTFQFEDCTVLTIAHRLNTIMDYTRVIL 1500  
 DB 1441 ARALLRTKILVLDEATAAVIDETDDLIQSTIRTFQFEDCTVLTIAHRLNTIMDYTRVIL 1500

QY 1501 DKGIQYVGSDDLQOGLFYSMAXDAGLV 1531  
 DB 1501 DKGIQYVGSDDLQOGLFYSMAXDAGLV 1531

RESULT 8  
 Q8HX05 PRELIMINARY; PRT; 1530 AA.  
 ID Q8HX05 PRELIMINARY; PRT; 1530 AA.  
 AC Q8HX05;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Multidrug resistance protein 1.  
 GN Name=MRP1;  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovinae; Bos.

NCBI\_TaxID=9913;  
 [1]  
 SEQUENCE FROM N.A.  
 MEDLINE=22062955; PubMed=12067707; DOI=10.1016/S0014-5793(02)02816-8;  
 Taguchi Y., Saki K., Konano T.;  
 RA "functional analysis of MRPI cloned from bovine."  
 RT FEBS Lett. 521:211-213(2002).  
 CC -!- SIMILARITY: Belongs to the ABC transporter family.  
 DR EMBL; AB082124; BAC1550.1; -;  
 DR HSSP; P08716; LMT0.  
 DR GO; GO:0016021; C: integral to membrane; IEA.  
 DR GO; GO:0005524; F: ATP binding; IEA.  
 DR GO; GO:0042626; F: ATPase activity, coupled to transmembrane m. . .; IEA.  
 DR GO; GO:0000166; F: nucleotide binding; IEA.  
 DR GO; GO:0005215; F: transporter activity; IEA.  
 DR GO; GO:0006810; P: transport; IEA.  
 DR InterPro; IPR003593; AAA\_ATPase.  
 DR InterPro; IPR011527; ABC membrane 1.  
 DR InterPro; IPR001140; ABC TM transp.  
 DR InterPro; IPR003439; ABC transporter.  
 DR InterPro; IPR005292; MRP\_assoc.  
 DR Pfam; PF00664; ABC membrane; 2.  
 DR Pfam; PF00005; ABC\_tran; 2.  
 DR ProDom; PD000006; ABC\_transporter; 2.  
 DR SMART; SM00382; AAA; 2.  
 DR TIGRFAM; TIGR00957; MRP\_assoc\_pro; 1.  
 DR PROSITE; PS00929; ABC\_TMIF; 2.  
 DR PROSITE; PS00211; ABC\_TRANSPORTER\_1; 2.  
 DR PROSITE; PS00893; ABC\_TRANSPORTER\_2; 2.  
 KW ATP-binding.  
 SQ SEQUENCE 1530 AA; 171664 MW; AAE4F92ED7832703 CRC64;  
  
 Query Match 73.0%; Score 7230.5; DB 2; Length 1530;  
 Best Local Similarity 90.7%; Pred. No. 0;  
 Matches 1388; Conservative 83; Mismatches 59; Indels 1; Gaps 1;  
  
 QY 1 MALRGFCSADGSDPLDMNVTWNTSNPDKFQNTLVWVPCFYLVWACFPFYLGRH 60  
 DB 1 MALRDFCSVDGSDLFWENVTWNTSNPDKFQNTLVWVPCFYLVWACFPFYLGRH 60  
  
 QY 61 DRGYQMTPLNKTALGFLMIVCWADLFYFWERSGIFLAPFLVSPITLLGTTLLA 120  
 DB 61 DRGYQMTPLNKTALGFLMIVCWADLFYFWERSGIFLAPFLVSPITLLGTTLLA 120  
  
 QY 121 TELIOLERRKGVOSSGIMTLFWLVALVCALALIRSKIMTALKEDAQVDLFRDITFVYVS 180  
 DB 121 TELIQIERRRGVOSSGIMTLFWLVALVCALALIRSKIMTALKEDARVDVFRDITFVYVS 180  
  
 QY 181 LLLIQLVLSGFCSDRSPFSETIHDNPPCPSSASFLSRITFWITGLIVRGYRQPLEGSD 240  
 DB 181 LVLIQLVLSGFCSDRSPFSETINDNPPCPSSASFLSRITFWITGMVQYRQPLESTD 240  
  
 QY 241 LMSLNKEDTSEQVVPVLVQWKECAKTRKQPVVYSSKDPAPKESKVDANEVEAL 300  
 DB 241 LMSLNKEDTSEQVVPVLVQWKECAKSKQPVVYSSKDPAPKESKVDNVEAEAL 300  
  
 QY 301 IVKSPQKWNPSLFVLYKTFGPFYLMSPFPAIHDLMMFSGPQILKLLIKFVNDTKAPD 360  
 DB 301 IVKQKQERDPSLFVLYKTFGPFYLMSPFLKAVHDLMMFAGPEILKLLINFVNDKKAPE 360  
  
 QY 361 WQGYFYTVLLFVYACQLTLVHLQYFHICFVSGMRKTAIVGAVRKALVITNARKSTV 420  
 DB 361 WQGYFYTVLLFVYACQLTLVHLQYFHICFVSGMRKTAIVGAVRKALVITNARKSTV 420  
  
 QY 421 GEIVNLSVDAQRFDLATYINNIWSAPLOVILALYLLWNLGSPVAGVAVMLVMPVN 480  
 DB 421 GEIVNLSVDAQRFDLATYINNIWSAPLOVILALYLLWNLGSPVAGVAVMLVMPVN 480  
  
 QY 481 AVAMKTKTYQVAHMKSKDNRIKLWNEILNGIKVLKLYAWELAFKDKVLAIQBELKVLK 540  
 DB 481 AVAMKTKTYQVAHMKSKDNRIKLWNEILNGIKVLKLYAWELAFKDKVLAIQBELKVLK 540  
  
 QY 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTTIDENNILDAQAFVSLALFNILRFPFLNLP 600

DB 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTTIDENNILDAQAFVSLALFNILRFPFLNLP 600  
 QY 601 MVISSIVQASVSLKELRIFLSHEELEPPDISERRPVKDCGGTNSITVRNATFTWASDPT 660  
 DB 601 MVISSIVQASVSLKELRIFLSHEELEPPDISERRPVKDCGGTNSITVRNATFTWASDPT 660  
 QY 661 LMGITFSIPGALVAVVQVCGKSSLSALLAEMDKVEGHVIAIKGSVAYVPPQAWIQND 720  
 DB 661 LMGITFSIPGALVAVVQVCGKSSLSALLAEMDKVEGHVIAIKGSVAYVPPQAWIQND 720  
 QY 721 SLRENILFGCOLBEPYRSVIOACALLPDLLEILPSGDRTEIGEGVNLSSGGQKQVSLAR 780  
 DB 721 SLRENILFGCOLBEPYRSVIOACALLPDLLEILPSGDRTEIGEGVNLSSGGQKQVSLAR 780  
 QY 781 AVYSNADILYLPDDPLSADVAHVGHIFENVIKPGKMLKNTKRLIVTHSMYSYLPQVDVIV 840  
 DB 781 AVYCDSDVILLDDPLSADVAHVGHIFENVIKPGKMLKNTKRLIVTHSMYSYLPQVDVIV 840  
 QY 841 MSGGKISEMGSYQELLARDGAFAPLRTYASTEQSDAEENGVTGSGPGKEAKOMENGM 900  
 DB 841 MSGGKISEMGSYQELLARDGAFAPLRTYASTEQSDAEENGVTGSGPGKEAKOMENGM 900  
 QY 901 LVTSAGKQLQRLSSSSSYSGDISRHNSHTAELOKAEAKKEETWKLMEADKAQTCQVKL 960  
 DB 901 LVTDTAGKMQRLSSSSSYSGDISRHNSHTAELOKAEAKKEETWKLMEADKAQTCQVKL 960  
 QY 961 SVYDYMKAIGLFTSFLSIFLPMCNHVSALASNYWLSLWTDPIVNGTQEHKTVRLSVYG 1020  
 DB 961 SVYDYMKAIGLFTSFLSIFLPMCNHVSALASNYWLSLWTDPIVNGTQEHKTVRLSVYG 1020  
 QY 1021 ALGISQIAGVFGYSMAVSIIGGILASRCLHVDLLHSILRSPMSFFERTSGNLVNRFSKEL 1080  
 DB 1020 ALGISQITVFGYSMAVSIIGGIFASRLHLDLHNLASPSIFFERTSGNLVNRFSKEL 1079  
 QY 1081 DTVDSMIPVIMPMGSLFNIVGACIVILLATPIAAIIIPPLGLIYFFVQRYVYASSROL 1140  
 DB 1081 DTVDSMIPVIMPMGSLFNIVGACIVILLATPIAAIIIPPLGLIYFFVQRYVYASSROL 1139  
 QY 1141 KRLESVSRSPVYSHNETLLGVSVIRAFEEQERFIHQSDLKVDENOKAYYPSIVANRWLA 1200  
 DB 1141 KRLESVSRSPVYSHNETLLGVSVIRAFEEQERFIHQSDLKVDENOKAYYPSIVANRWLA 1199  
 QY 1201 VRLECVGNICVILFALFPAVISRHSLSAGLVLSVSYSLQVTTYLNLVMSSEMETNIYA 1260  
 DB 1200 VRLECVGNICVILFALFPAVISRHSLSAGLVLSVSYSLQVTTYLNLVMSSEMETNIYA 1259  
 QY 1261 VERLKEYSETEKEAPWQIQETAPPSSWPQVGRVFNRYCLYREDLDLDFVLRHINTINGG 1320  
 DB 1260 VERLKEYSETEKEAPWQIQETAPPSSWPQVGRVFNRYCLYREDLDLDFVLRHINTINGG 1319  
 QY 1321 EKVGVGTGTGAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPVLP 1380  
 DB 1320 EKVGVGTGTGAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPVLP 1379  
 QY 1381 SGSLRMNLDPSQSDSEEVWTSLELAHLKDFVSALPDKLDHECAEGENLSVGQRLVCL 1440  
 DB 1380 SGSLRMNLDPSQSDSEEVWTSLELAHLKDFVSALPDKLDHECAEGENLSVGQRLVCL 1439  
 QY 1441 ARALLRKTILVLDATAVLETDLLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 DB 1440 ARALLRKTILVLDATAVLETDLLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1499  
 QY 1501 DKGETQEGAPSDLLQQRGLFYSMADAGLV 1531  
 DB 1500 DKGETQEGAPSDLLQQRGLFYSMADAGLV 1530

RESULT 9  
 Q9UQ98 PRELIMINARY; PRT; 1400 AA.  
 AC Q9UQ98;  
 DT 01-MAY-2000 (Tremblrel. 13, Created)

DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)	Db	1	DWNTWNTSNDFTKCFQNTVLVWVPCFYLNACFPFFLYLSRHRGVIQMTPLNKTXTA	60
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)	Qy	77	LGFLWIVCWADLFYSPWERSRGIFLAPVFLVSPTLGLITTLATFLTLQERRRGVQSSG	136
DE	Multidrug resistance protein (Fragment).	Db	61	LGFLWIVCWADLFYSPWERSRGIFLAPVFLVSPTLGLITTLATFLTLQERRRGVQSSG	120
GN	Name=MRP;	Qy	137	IMLFWLVALVCAALAILRSKIMTALKEDAQVDLFRDITFYVYFSLILLIQLVLSFSDRSP	196
OS	Homo sapiens (Human).	Db	121	IMLFWLVALVCAALAILRSKIMTALKEDAQVDLFRDITFYVYFSLILLIQLVLSFSDRSP	180
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	Qy	197	LFSETHDNPNCPESSASFLSRITFWITGLIVRGYQPLEGSDLWSLNKEDTSEQVVPV	256
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.	Db	181	LFSETHDNPNCPESSASFLSRITFWITGLIVRGYQPLEGSDLWSLNKEDTSEQVVPV	240
OX	NCBI_TaxID=9606;	Qy	257	LVKWKKCEAKTRKOPVKVYSSKDPAPQKESKVDANEVEEALIVKSPQKWNPSLPKV	316
RV	SEQUENCE FROM N.A.	Db	241	LVKWKKCEAKTRKOPVKVYSSKDPAPQKESKVDANEVEEALIVKSPQKWNPSLPKV	300
RP	MEDLINE=98008927; PubMed=9344662; DOI=10.1006/geno.1997.4950;	Qy	317	LYKTFGPFYFLMSPFFKAHDLMMESGPQILKLLIKFVNDTKAPDQOGYFTVLLFVTA	376
RX	Grant C.E., Kurz E.U., Cole S.P., Deeley R.G.;	Db	301	LYKTFGPFYFLMSPFFKAHDLMMESGPQILKLLIKFVNDTKAPDQOGYFTVLLFVTA	360
RA	"Analysis of the intron-exon organization of the human multidrug-	Qy	377	QTLVLHQYFHCYFVSGMRITKAVIGAVYRKALVITNSARKSSTVGEIVNLSVDAQRFMD	436
RE	resistance protein gene (MRP) and alternative splicing of its mRNA.";	Db	361	QTLVLHQYFHCYFVSGMRITKAVIGAVYRKALVITNSARKSSTVGEIVNLSVDAQRFMD	420
RL	Genomics 45:368-378(1997).	Qy	437	LATVINMWSAPLOVILALYLLWNLGSPVLGAVVMVLPVNVAMAMTKTYQVAHMK	496
CC	-1- SIMILARITY: belongs to the ABC transporter family.	Db	421	LATVINMWSAPLOVILALYLLWNLGSPVLGAVVMVLPVNVAMAMTKTYQVAHMK	480
DR	EMBL; AF022824; AAB83982.1; JOINED.	Qy	497	SKDNRIKLMNEILNGIKVLKLYAWELAFKDKVLAIROBELKVLKKSAYLSAVGTFTWVCT	556
DR	EMBL; AF022825; AAB83982.1; JOINED.	Db	481	SKDNRIKLMNEILNGIKVLKLYAWELAFKDKVLAIROBELKVLKKSAYLSAVGTFTWVCT	540
DR	EMBL; AF022826; AAB83982.1; JOINED.	Qy	557	PFLVALCTFAVYVTTIDENNILDAQTAQFVSLAFNLFPFLNIPMWISSIVQASVSKRL	616
DR	EMBL; AF022827; AAB83982.1; JOINED.	Db	541	PFLVALCTFAVYVTTIDENNILDAQTAQFVSLAFNLFPFLNIPMWISSIVQASVSKRL	600
DR	EMBL; AF022828; AAB83982.1; JOINED.	Qy	617	RIFLSHELEPDSIERRPVKGGTNSITVRNATFTWARSOPPTLNGITITSIPGALVAV	676
DR	EMBL; AF022829; AAB83982.1; JOINED.	Db	601	RIFLSHELEPDSIERRPVKGGTNSITVRNATFTWARSOPPTLNGITITSIPGALVAV	660
DR	EMBL; AF022830; AAB83982.1; JOINED.	Qy	677	VGQVCGKSSLLSALLAEMDKVEGHVAIKGSVAVPOQAWIQNDLSRENILFCQLEPEY	736
DR	EMBL; AF022831; AAB83982.1; JOINED.	Db	661	VGQVCGKSSLLSALLAEMDKVEGHVAIKGSVAVPOQAWIQNDLSRENILFCQLEPEY	688
DR	EMBL; AF022832; AAB83982.1; JOINED.	Qy	737	YRSVIOACALLPDLLEILPSGDRTEIGBKGNLGGQKQVSLARAVYSNADYILFDDPLS	796
DR	EMBL; AF022833; AAB83982.1; JOINED.	Db	689	YRSVIOACALLPDLLEILPSGDRTEIGBKGNLGGQKQVSLARAVYSNADYILFDDPLS	688
DR	EMBL; AF022834; AAB83982.1; JOINED.	Qy	797	AVDAHVGKHI FENVIGPKGMLKNKTRILVTHSMYLPQVDVIVMSGGKI SEMGSIQELL	856
DR	EMBL; AF022835; AAB83982.1; JOINED.	Db	689	AVDAHVGKHI FENVIGPKGMLKNKTRILVTHSMYLPQVDVIVMSGGKI SEMGSIQELL	725
DR	EMBL; AF022836; AAB83982.1; JOINED.	Qy	857	ARDGAFAEFLRTYASTEQEQAEBENGVTGVSFGKEAKOMENGLVTDSDAGKQLQRLSS	916
DR	EMBL; AF022837; AAB83982.1; JOINED.	Db	726	ARDGAFAEFLRTYASTEQEQAEBENGVTGVSFGKEAKOMENGLVTDSDAGKQLQRLSS	785
DR	EMBL; AF022838; AAB83982.1; JOINED.	Qy	917	SSSYSGDISRRHNSAELOKAEAKKETWKLMEADKAQTCQVKLSVYWDYMKAIGLFISF	976
DR	EMBL; AF022839; AAB83982.1; JOINED.	Db	786	SSSYSGDISRRHNSAELOKAEAKKETWKLMEADKAQTCQVKLSVYWDYMKAIGLFISF	845
DR	EMBL; AF022840; AAB83982.1; JOINED.	Qy	977	LSIFLFCNHNVSALASNYLWLTDDPIVNGTQEHKTVRLSVYCNALGISOGIAVFGYSMA	1036
DR	EMBL; AF022841; AAB83982.1; JOINED.	Db	846	LSIFLFCNHNVSALASNYLWLTDDPIVNGTQEHKTVRLSVYCNALGISOGIAVFGYSMA	905
DR	EMBL; AF022842; AAB83982.1; JOINED.	Qy	1037	VSIGGILASRCLHVDLLHSILRSPMFFERTPSGNLVNRFSEKELDTVDMSIPEIKFMFG	1096
DR	EMBL; AF022843; AAB83982.1; JOINED.	Db	906	VSIGGILASRCLHVDLLHSILRSPMFFERTPSGNLVNRFSEKELDTVDMSIPEIKFMFG	965
DR	EMBL; AF022844; AAB83982.1; JOINED.	Qy	1097	SLFNVI GACIVILLATPIAAIIIPPLGLIYFFVQRFVAVSSRQLKRLSVSRPVSHPN	1156
DR	EMBL; AF022845; AAB83982.1; JOINED.	Db	966	SLFNVI GACIVILLATPIAAIIIPPLGLIYFFVQRFVAVSSRQLKRLSVSRPVSHPN	1025

Query Match 71.7%; Score 7104.5; DB 2; Length 1400;  
Best Local Similarity 92.4%; Pred. No. 0;  
Matches 1400; Conservative 0; Mismatches 0; Indels 115; Gaps 1;  
17 DWNTWNTSNDFTKCFQNTVLVWVPCFYLNACFPFFLYLSRHRGVIQMTPLNKTXTA 76  
|||||

QY 1157 ETLGVSIVRAFEQEERFIHQSDLKVDENOKAYPSIVANRWLVRLECVGNCIVLFAAL 1216  
 |||||  
 DB 1026 ETLGVSIVRAFEQEERFIHQSDLKVDENOKAYPSIVANRWLVRLECVGNCIVLFAAL 1085  
 |||||  
 QY 1217 FAVISRHSLSAGLVGSYSYQVTTYLNWLRMSSEMETNIVAVERLKEYSETEKEAPW 1276  
 |||||  
 DB 1086 FAVISRHSLSAGLVGSYSYQVTTYLNWLRMSSEMETNIVAVERLKEYSETEKEAPW 1145  
 |||||  
 QY 1277 QIQETAPPSWQVQVEPRNYCLRYREDLDFVLRHINTVINGEKVGIVGRTGAGKSSL 1336  
 |||||  
 DB 1146 QIQETAPPSWQVQVEPRNYCLRYREDLDFVLRHINTVINGEKVGIVGRTGAGKSSL 1205  
 |||||  
 QY 1337 TLGLFRINSEAGEIIGDINIAGLHDLRFKTIIPQDPVLFSGSLRMNLDPPFSQYSD 1396  
 |||||  
 DB 1206 TLGLFRINSEAGEIIGDINIAGLHDLRFKTIIPQDPVLFSGSLRMNLDPPFSQYSD 1265  
 |||||  
 QY 1397 EEWTSLELAHLKDFVSALPKDLHECAEGGNSVGQRLVCLARALLKTKILVLDEA 1456  
 |||||  
 DB 1266 EEWTSLELAHLKDFVSALPKDLHECAEGGNSVGQRLVCLARALLKTKILVLDEA 1325  
 |||||  
 QY 1457 TAAVLETDLLQSIQRTQFECTVLTIAHRLNTIMDYTRVIVLDKGEIOEYGAESDILQ 1516  
 |||||  
 DB 1326 TAAVLETDLLQSIQRTQFECTVLTIAHRLNTIMDYTRVIVLDKGEIOEYGAESDILQ 1385  
 |||||  
 QY 1517 QRGLFYMAKADAGLV 1531  
 |||||  
 DB 1386 QRGLFYMAKADAGLV 1400  
 |||||

RESULT 10  
 MRPI\_MOUSE  
 ID MRPI\_MOUSE STANDARD; PRT; 1528 AA.  
 AC Q35379;  
 DT 25-OCT-2004 (Rel. 45, Created)  
 DT 25-OCT-2004 (Rel. 45, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Multidrug resistance-associated protein 1 (ATP-binding cassette, sub-family C, member 1).  
 GN Name=Abcc1; Synonyms=Abcc1a, Abcc1b, Mdrap, Mip;  
 OS Mus musculus (Mouse)  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxId=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Muscle;  
 RX MEDLINE=96251691; PubMed=8649356;  
 RA Stride B.D., Valdimarsson G., Gerlach J.H., Wilson G.M., Cole S.P., Deeley R.G.;  
 RT "Structure and expression of the messenger RNA encoding the murine multidrug resistance protein, an ATP-binding cassette transporter.";  
 RL Mol. Pharmacol. 49:962-971(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Testis;  
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;  
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S., Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H., Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T., Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J., Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W., Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S., Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S., Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J., Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D., Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L., Koyama A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A., Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H., Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G., Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S., Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M., Sandelin A., Schneider C., Sempke C.A., Setou M., Shimada K.,

Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M., Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C., Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L., Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N., Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K., Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S., Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa A., Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A., Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J., Birney E., Hayashizaki Y.;  
 "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.";  
 RL Nature 420:563-573(2002).  
 RT  
 CC -!- FUNCTION: May participate directly in the active transport of drugs into subcellular organelles or influence drug distribution indirectly (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).  
 CC -!- SIMILARITY: Belongs to the ABC transporter family. MRP subfamily.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 CC EMBL; AF022908; AAB80938.1; -;  
 CC EMBL; AK029876; BAC26654.1; -;  
 CC HSP; F08716; IMT0.  
 CC MGD; MGI:102676; Abcc1.  
 CC InterPro; IPR003593; AAA\_ATPase.  
 CC InterPro; IPR011527; ABC\_membrane\_1.  
 CC InterPro; IPR011140; ABC\_TM\_transp.  
 CC InterPro; IPR003439; ABC\_transporter.  
 CC InterPro; IPR005292; MRP\_assoc.  
 CC Pfam; PF00664; ABC\_membrane; 2.  
 CC Pfam; PF00005; ABC\_tran; 2.  
 CC ProDom; PD000006; ABC\_transporter; 2.  
 CC SMART; SM00382; AAA; 2.  
 CC TIGRFAMs; TIGR00957; MRP\_assoc\_pro; 1.  
 CC PROSITE; PS50929; ABC\_TMIF; 2.  
 CC PROSITE; PS00211; ABC\_TRANSPORTER\_1; 2.  
 CC PROSITE; PS50893; ABC\_TRANSPORTER\_2; 2.  
 KW ATP-binding; Glycoprotein; Repeat; Transmembrane; Transport.  
 FT DOMAIN 1 33 Extracellular (By similarity).  
 FT TRANSMEM 34 54 1 (By similarity).  
 FT DOMAIN 55 74 Cytoplasmic (By similarity).  
 FT TRANSMEM 75 95 2 (By similarity).  
 FT DOMAIN 96 100 Extracellular (By similarity).  
 FT TRANSMEM 101 121 3 (By similarity).  
 FT DOMAIN 122 133 Cytoplasmic (By similarity).  
 FT TRANSMEM 134 154 4 (By similarity).  
 FT DOMAIN 155 172 Extracellular (By similarity).  
 FT TRANSMEM 173 193 5 (By similarity).  
 FT DOMAIN 194 317 Cytoplasmic (By similarity).  
 FT TRANSMEM 318 338 6 (By similarity).  
 FT DOMAIN 339 364 Extracellular (By similarity).  
 FT TRANSMEM 365 385 7 (By similarity).  
 FT DOMAIN 386 441 Cytoplasmic (By similarity).  
 FT TRANSMEM 442 462 8 (By similarity).  
 FT DOMAIN 463 465 Extracellular (By similarity).  
 FT TRANSMEM 466 486 9 (By similarity).  
 FT DOMAIN 487 548 Cytoplasmic (By similarity).  
 FT TRANSMEM 549 569 10 (By similarity).  
 FT DOMAIN 570 591 Extracellular (By similarity).  
 FT TRANSMEM 592 612 11 (By similarity).  
 FT DOMAIN 613 963 Cytoplasmic (By similarity).  
 FT TRANSMEM 964 984 12 (By similarity).  
 FT DOMAIN 985 1022 Extracellular (By similarity).  
 FT TRANSMEM 1023 1043 13 (By similarity).  
 FT DOMAIN 1044 1086 Cytoplasmic (By similarity).  
 FT TRANSMEM 1087 1107 14 (By similarity).

FT	DOMAIN	1108	1108	Extracellular (By similarity).	
FT	TRANSMEM	1109	1129	15 (By similarity).	
FT	DOMAIN	1130	1200	Cytoplasmic (By similarity).	
FT	TRANSMEM	1201	1221	16 (By similarity).	
FT	DOMAIN	1222	1233	Extracellular (By similarity).	
FT	TRANSMEM	1234	1244	17 (By similarity).	
FT	DOMAIN	1245	1528	Cytoplasmic (By similarity).	
FT	DOMAIN	1246	868	ABC transporter 1.	
FT	DOMAIN	1250	1524	ABC transporter 2.	
FT	NP_BIND	678	695	ATP (Potential).	
FT	NP_BIND	1324	1331	ATP (Potential).	
FT	CARBOHYD	19	19	N-linked (GlcNAc. . .) (Potential).	
FT	CARBOHYD	1003	1003	N-linked (GlcNAc. . .) (Potential).	
SQ	SEQUENCE	1528 AA;	171183 MW;	68FD13667D61DBBB CRC64;	
Query Match					
Best Local Similarity 88.0%; Pred. No. 0;					
Matches 1349; Conservative 102; Mismatches 75; Indels 7; Gaps 5;					
QY	1	MALRFGCSADGSDPLMDWNTWNTGNPTKCFQNTLVWVPCFYLMACFPYFLYLSRH	60		
DB	1	MALRFGCSADGSDPLMDWNTWNTGNPTKCFQNTLVWVPCFYLMACFPYFLYLSRH	60		
QY	61	DRGYTQMTPLNKTALGFLMIVCWADLFYSFWERSRGIFLAPFLVSPITLLGTTLLA	120		
DB	61	DRGYTQMTPLNKTALGFLMIVCWADLFYSFWERSRGIFLAPFLVSPITLLGTTLLA	120		
QY	121	TLFIOLERRKGVOSSGIMLFWLWALCALAILRSKIMTALKEDAOVDFRDIIFYVYFS	180		
DB	121	TLFIOLERRKGVOSSGIMLFWLWALCALAILRSKIMTALKEDAOVDFRDIIFYVYFS	180		
QY	181	LLLIQLVLSCFSDRPLFSETHDPNCPSSASFLSRITFWITGLVGRQPLEGSD	240		
DB	181	LVLQVLSCFSDRPLFSETHDPNCPSSASFLSRITFWITGLVGRQPLEGSD	240		
QY	241	LWSLNKEDTSQVVPVLVKNWKECAKTRKQPKVYVS-SKDPAPQKSSSKVDANEEVEA	299		
DB	241	LWSLNKEDTSQVVPVLVKNWKECAKTRKQPKVYVS-SKDPAPQKSSSKVDANEEVEA	299		
QY	300	LIVKSPQKEMPSLFKVLKTFPGFVFLMSFPFKAJHDLMMFSGPOLKLLIKFVNDTKAP	359		
DB	301	LIVKSPHNDRPSLFKVLKTFPGFVFLMSFPFKAJHDLMMFSGPOLKLLIKFVNDTKAP	360		
QY	360	DWQGYFTVLLFVTAQTLVHOFHICFVSGMRITKAVIGAVYRKALVITNSARKST	419		
DB	361	DWQGYFTVLLFVTAQTLVHOFHICFVSGMRITKAVIGAVYRKALVITNSARKST	420		
QY	420	VGEIVNLSVDAQRPMDLATYINMIWSAPLOVILALYLLMLNLSGLAGVAVMLVMPV	479		
DB	421	VGEIVNLSVDAQRPMDLATYINMIWSAPLOVILALYLLMLNLSGLAGVAVMLVMPV	480		
QY	480	NAVMAWTKTKYQVAHMKSKDNRIKLMMELNGIKVLKYAWELAPQDKVLAIRQEEKVL	539		
DB	481	NAVMAWTKTKYQVAHMKSKDNRIKLMMELNGIKVLKYAWELAPQDKVLAIRQEEKVL	540		
QY	540	KKSAYLSAVGFTTWCTPFLVALCTFAVYVVIDENNILDAQTAFLVSLALFNILAPFNIL	599		
DB	541	KKSAYLSAVGFTTWCTPFLVALCTFAVYVVIDENNILDAQTAFLVSLALFNILAPFNIL	600		
QY	600	PMWISSIVQASVSLKRLIFISHELEPDSTERRPVKDGGTNSITVKNATFTWARSPPP	659		
DB	601	PMWISSIVQASVSLKRLIFISHELEPDSTERRPVKDGGTNSITVKNATFTWARSPPP	659		
QY	660	TLNGITTFSPREGALVAVGVQCGKSLLSALLAEMDKVEGHVAKGSVAVVPOQAWTON	719		
DB	660	TLNGITTFSPREGALVAVGVQCGKSLLSALLAEMDKVEGHVAKGSVAVVPOQAWTON	719		
QY	720	DSLRENILFGQLEPPYRSVIOACALLPDLEILPSGDRTEIGKGNVLSGQQRVSLSA	779		
DB	720	DSLRENILFGQLEPPYRSVIOACALLPDLEILPSGDRTEIGKGNVLSGQQRVSLSA	779		
QY	780	RAVYNSADIYLFDDPLSAVDHVGKHPENVIGPKGMLKNKTRILVTHSMYSYLPOVDVII	839		

DB	780	RAVYNSADIYLFDDPLSAVDHVGKHPENVIGPKGMLKNKTRILVTHSMYSYLPOVDVII	839		
QY	840	VMSGKISEMSYQELLARDGAFAPFIRTYASTEQEODAEENGVTGCVSGPKAEKQMGNG	899		
DB	840	VMSGKISEMSYQELLARDGAFAPFIRTYASTEQEODAEENGVTGCVSGPKAEKQMGNG	899		
QY	900	MLVTDGAKQORQLSSSSSSGDISSRHNTSTAELOKAEKKEETWKLMEADKATGOVK	959		
DB	897	MLVTDGAKQORQLSSSSSSGDISSRHNTSTAELOKAEKKEETWKLMEADKATGOVK	955		
QY	960	LSVYDYMKAIGLPIFLSIFLFCMNVHVSALASNYLSLWTD--PIVNGTQSHTKVRLSV	1018		
DB	956	LSVYDYMKAIGLPIFLSIFLFCMNVHVSALASNYLSLWTD--PIVNGTQSHTKVRLSV	1015		
QY	1019	YGALGISOGIAVFGYSMAVSGIGILASRCIHDVLLHSLRSPMSFFETPSGNLVNRFSK	1078		
DB	1016	YGALGISOGIAVFGYSMAVSGIGILASRCIHDVLLHSLRSPMSFFETPSGNLVNRFSK	1075		
QY	1079	ELDTVDSDMPEVIXMFGSLFNVIACIVILLATPIAAIIIPPLGLIYFFVQRFVASSR	1138		
DB	1076	ELDTVDSDMPEVIXMFGSLFNVIACIVILLATPIAAIIIPPLGLIYFFVQRFVASSR	1135		
QY	1139	QLKRLSVSRSPVYSHFNETLLGVSVIRAFPEQERFIHQSDLKVDENQKAYYPSIVANRW	1198		
DB	1136	QLKRLSVSRSPVYSHFNETLLGVSVIRAFPEQERFIHQSDLKVDENQKAYYPSIVANRW	1195		
QY	1199	LAVLECVGNCIVLFAALFAVISRHSLSAGLVGSVSLQVTTVYLNKLVMSSEMETNI	1258		
DB	1196	LAVLECVGNCIVLFAALFAVISRHSLSAGLVGSVSLQVTTVYLNKLVMSSEMETNI	1255		
QY	1259	VAVERLKEYSTETKEAPMIOETAPPSPQVGRVFRNYCLRYREDLDFVLRHINVTIN	1318		
DB	1256	VAVERLKEYSTETKEAPMIOETAPPSPQVGRVFRNYCLRYREDLDFVLRHINVTIN	1315		
QY	1319	GGEKGVIGRTGAKSSLTGLFRINSAEGLIIDGINIAKIHLHDLRFKTIIPQDPV	1378		
DB	1316	GGEKGVIGRTGAKSSLTGLFRINSAEGLIIDGINIAKIHLHDLRFKTIIPQDPV	1375		
QY	1379	LFPSGLRNLDPPFSQSDSEEVWTSLELAHLKDFYSALPDKLHCEAEGENLSVGOROLV	1438		
DB	1376	LFPSGLRNLDPPFSQSDSEEVWTSLELAHLKDFYSALPDKLHCEAEGENLSVGOROLV	1435		
QY	1439	CLARALLRKTKILVLDLATAVLETDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVI	1498		
DB	1436	CLARALLRKTKILVLDLATAVLETDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVI	1495		
QY	1499	VLDKGEVQEGAPSDLLQORGLFYSMADAGLV	1531		
DB	1496	VLDKGEVQEGAPSDLLQORGLFYSMADAGLV	1528		
RESULT 11					
ID	Q810E4	PRELIMINARY;	PRT;	1532 AA.	
AC	Q810E4;				
DT	01-JUN-2003	(TrEMBLrel. 24, Created)			
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)			
DT	01-MAR-2004	(TrEMBLrel. 26, Last annotation update)			
DE	ATP-binding cassette protein Cl.				
OS	Rattus norvegicus (Rat).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.				
OX	NCBI_TaxID=10116;				
EN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=Sprague-Dawley; TISSUE=Spleen;				
RA	Yabuuchi H., Takayanagi S., Ishikawa T.;				
RL	Submitted (PDB-2002) to the EMBL/GenBank/DBJ databases.				
CC	-1- SIMILARITY: Belongs to the ABC transporter family.				
DR	EMBL; AF487549; AA085437.1; -				
DR	HSSP; P08716; IMT0.				
DR	GO; GO:0016021; C:integral to membrane; IEA.				
DR	GO; GO:0005524; F:ATP binding; IEA.				

DR GO: 0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.  
 DR GO: 0000166; F:nucleotide binding; IEA.  
 DR GO: 0005215; F:transporter activity; IEA.  
 DR GO: 0006810; P:transport; IEA.  
 DR InterPro: IPR003593; AAA\_ATPase.  
 DR InterPro: IPR011527; ABC\_membrane\_1.  
 DR InterPro: IPR001140; ABC\_TM\_transpt.  
 DR InterPro: IPR003439; ABC\_transporter.  
 DR InterPro: IPR001395; Aldo/ket\_red.  
 DR InterPro: IPR005292; MRP\_assoc.  
 DR Pfam: PF00664; ABC\_membrane\_2.  
 DR Pfam: PF00005; ABC\_tran; 2.  
 DR ProDom: PD00006; ABC\_transporter; 2.  
 DR SMART: SM00382; AAA; 2.  
 DR TIGRFAMs: TIGR00957; MRP\_assoc\_pro; 1.  
 DR PROSITE: PS00211; ABC\_TMIF; 2.  
 DR PROSITE: PS00211; ABC\_TRANSPORTER\_1; 2.  
 DR PROSITE: PS00893; ABC\_TRANSPORTER\_2; 2.  
 DR PROSITE: PS00063; ALDO\_KETO\_REDUCTASE\_3; UNKNOWN\_1.  
 KW ATP-binding.  
 SQ SEQUENCE 1532 AA; 171491 MW; 2E6939F63F5A3F68 CRC64;

Query Match 70.0%; Score 6932.5; DB 2; Length 1532;  
 Best Local Similarity 87.3%; Pred. No. 0;  
 Matches 1339; Conservative 99; Mismatches 92; Indels 3; Gaps 3;

QY 1 MALRGCADGSDPLMDWNTWNTGNPDFTKCFQNTLVWVPCFYLWACFPYFLYLSRH 60  
 DB 1 MALSFCSGSDPLMDWNTWNTGNPDFTKCFQNTLVWVPCFYLWACFPYFLYLSRH 60

QY 61 DRGYTOMTLNKTALGFLMLVWADLFYSFWERSRGI FLAPVFLVSP TLLGTTLLA 120  
 DB 61 DRGYTOMTLNKTALGFLMLVWADLFYSFWERSRGI FLAPVFLVSP TLLGTTLLA 120

QY 121 TFLIQLERRKGVOSSGIMLTFWLVALCALILRSKIMTALKEDAQVDLPDITPYVYFS 180  
 DB 121 TFLIQLERRKGVOSSGIMLTFWLVALCALILRSKIMTALKEDAQVDLPDITPYVYFS 180

QY 181 LILLIQLVSCFDRPLFSETTHDNPCESSASFLSRITFWITGLIVRGYRQPLEGSD 240  
 DB 181 LVFIQLVSCFDSPLFSETVDRNPCESSASFLSRITFWITGLIVRGYRQPLEGSD 240

QY 241 LWSLNKEDTSEQVPLVKNWKECAKTRKQPKVYVS -SKDPAQPKESSKVDANEEVEA 299  
 DB 241 LWSLNKEDTSEQVPLVKNWKECAKTRKQPKVYVS -SKDPAQPKESSKVDANEEVEA 299

QY 300 LIVKSPQKEWNPSPFLKVLKTPGPFYLSFPFKALHDLMMSPGPOILKLLIKFVNDTKAP 359  
 DB 301 LIVKSSHKDRDPSLFKVDYKTFGPFYLSFYLKALHDLMMFAGPPEILELIINFVNDREAP 360

QY 360 DWQGYFYTVLLFVTACLOTVLHQLYFHICFVSGMRIKTA VIGAVYRKALVITNSARKSST 419  
 DB 361 DWQGYLYTALLFVSACTOTLALHQLYFHICFVSGMRIKTA VIGAVYRKALVITNSARKSST 420

QY 420 VGEIVNLMSVDAQRFMDLATYINMTWSAPLQVILALYLWNLGSPVLGAVVWLMVFPV 479  
 DB 421 VGEIVNLMSVDAQRFMDLATYINMTWSAPLQVILALYLWNLGSPVLGAVVWLMVFPV 480

QY 480 NAWMAKTKTYQVAHMKSKDNRIKLWNLINGIKVLKYLAELAFKDKVLAIRQBELKVL 539  
 DB 481 NAWMAKTKTYQVAHMKSKDNRIKLWNLINGIKVLKYLAELAFKDKVLAIRQBELKVL 540

QY 540 KKSAYLSAVGTFTWCTPPLVALCTFAVYVTIDENNLDAOTAFVSLAFNLIRPLNIL 599  
 DB 541 KKSAYLAAGTFTWCTPPLVALSTFAVYVTIDENNLDAOTAFVSLAFNLIRPLNIL 600

QY 600 PMVISSIVQASVSLKRLRIFLSHEELEPDSTERRPVDGGGTNSITVRNATFTWARSDDP 659  
 DB 601 PMVISSIVQASVSLKRLRIFLSHEELEPDSTERRPVDGGGTNSITVRNATFTWARSDDP 660

QY 660 TLNGITFSPGALVAVGVQCGCKSSLLSALLAEMDKVEGHVALKGSVAVYVQQAQWQ 719  
 DB 661 TLNGITFAIPDGA VAVGVQCGCKSSLLSALLAEMDKVEGHVALKGSVAVYVQQAQWQ 720

QY 720 DSLRENILFGQQLBEPYRYSVIAQCALLPDLLEILPSGDRTEIGERGVNLGGQKQVSLA 779  
 DB 721 DSLRENILFGRPLQEHCHYKAVMEACALLPDLLEILPSGDLTEIGERGVNLGGQKQVSLA 780

QY 780 RAVYSNADIYLPDDPLSAVDHVGKHIPENVIKPGMLKKNKTRILVTHSMVSLPOVDVII 839  
 DB 781 RAVYCNSDIYLLDDPLSAVDHVGKHIFKVPKMPGLLKNKTRILVTHGISTYLPQVDVII 840

QY 840 VMSGKISEMGSYOELLARDGAFAEFLRTYASTEOQDAEENGVTGVSQPGKEAKQMENG 899  
 DB 841 VMSGKISEMGSYOELLARDGAFAEFLRTYASTEOQDAEENGVTGVSQPGKEAKQMENG 900

QY 900 MLVTDGAKQLQRLSSSSSYSGDISRHNSHTAEQLKAEAKKEETWKLMEADKAQTGVYK 959  
 DB 901 ILVTDGAKQLQRLSSSSSYSGDISRHNSHTAEQLKAEAKKEETWKLMEADKAQTGVYK 959

QY 960 LSVYWDYKAIKGLFISFLSIFLPMCNHVSALASNYWLSLWTD -PIVNGTQEHKTVRLSV 1018  
 DB 960 LSVYWNMYKAIKGLCISFLSIFLPMCNHVSALASNYWLSLWTD -PIVNGTQEHKTVRLSV 1019

QY 1019 YGALGISOGIAVFGYSMAVISIGGILASRCLHVDLLHSLRSPMSFFERTPSGNLVNRFSK 1078  
 DB 1020 YGALGILQGVAVFGYSMAVISIGGIFASRLHLDLLQNVLRSPMSFFERTPSGNLVNRFSK 1079

QY 1079 ELDTVDSMIPVIRKMFMSLNFVIGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSR 1138  
 DB 1080 ELDTVDSMIPVIRKMFMSLNFVIGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSR 1139

QY 1139 QLKRLSVSRSPVSHFNETHLLGVSVIRAFBEQERFIHQSDLKVDENOKAYYPSIVANRW 1198  
 DB 1140 QLKRLSVSRSPVSHFNETHLLGVSVIRAFBEQERFIHQSDLKVDENOKAYYPSIVANRW 1199

QY 1199 LAVRLECGNCIVLFAALFAVIRSHLSAGLVGVSYSLOVTTVTLNWLVRMSSMETNI 1258  
 DB 1200 LAVRLECGNCIVLFAALFAVIRSHLSAGLVGVSYSLOVTTVTLNWLVRMSSMETNI 1259

QY 1259 VAVRLEKSEYSETEKEAPWQIOETAPPSPQVGRVEFRNYCLRYREDLDFVLRHINVTIN 1318  
 DB 1260 VAVRLEKSEYSETEKEAPWQIOETAPPSPQVGRVEFRNYCLRYREDLDFVLRHINVTIN 1319

QY 1319 GGEKVGIVGRTGAGKSSITLGLFRINESAGEEIIIDGINIAKIGLHDLRFKLTIIIPQDPV 1378  
 DB 1320 GGEKVGIVGRTGAGKSSITLGLFRINESAGEEIIIDGINIAKIGLHDLRFKLTIIIPQDPV 1379

QY 1379 LFSGLRNLDPPSOYSDDEEVMTSLELAHLKDFVSALPKLDHCEAEGENLSVGQRLV 1438  
 DB 1380 LFSGLRNLDPPSOYSDDEEVMTSLELAHLKDFVSALPKLDHCEAEGENLSVGQRLV 1439

QY 1439 CLARALLRKTILVLDEATAAVDLETDLLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVI 1498  
 DB 1440 CLARALLRKTILVLDEATAAVDLETDLLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVI 1499

QY 1499 VLDKGEIOEYGA PSDLLOOGLFYFSMAKDAGLV 1531  
 DB 1500 VLDKGEIOEYGA PSDLLOOGLFYFSMAKDAGLV 1532

RESULT 12  
 Q8CG09 PRELIMINARY; PRT; 1532 AA.  
 AC Q8CG09;  
 DT 01-MAR-2003 (TReMBLrel. 23, Created)  
 DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)  
 DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)  
 DE Multidrug resistance-associated protein 1.  
 GN Name=Mrp1;  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC	STRAIN=Sprague-Dawley; TISSUE=Brain;	
RA	Yang Z., Li C.S.W., Shen D.D., Ho R.J.Y.;	
RL	Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.	
CC	1- SIMILARITY: Belongs to the ABC transporter family.	
DR	EMBL; AL170916; AN86532.1; -	
DR	HSP; P08716; IMT0.	
DR	GO; GO:0016021; C: integral to membrane; IEA.	
DR	GO; GO:0005524; F: ATP binding; IEA.	
DR	GO; GO:0043262; F: ATPase activity, coupled to transmembrane m. . . ; IEA.	
DR	GO; GO:0001666; F: nucleotide binding; IEA.	
DR	GO; GO:0005215; F: transporter activity; IEA.	
DR	GO; GO:0008610; P: transport; IEA.	
DR	InterPro; IPR003593; AAA ATPase.	
DR	InterPro; IPR011527; ABC membrane 1.	
DR	InterPro; IPR011140; ABC TM transp.	
DR	InterPro; IPR003439; ABC transporter.	
DR	InterPro; IPR001395; ABC/ket_red.	
DR	InterPro; IPR005292; MRP assoc.	
DR	Pfam; PF00664; ABC membrane; 2.	
DR	Pfam; PF00005; ABC tran; 2.	
DR	ProDom; PD000006; ABC transporter; 2.	
DR	SMART; SM00382; AAA; 2.	
DR	TIGRFAMs; TIGR00957; MRP assoc_pro; 1.	
DR	PROSITE; PS00929; ABC TMIF; 2.	
DR	PROSITE; PS00211; ABC TRANSPORTER_1; 2.	
DR	PROSITE; PS00893; ABC TRANSPORTER_2; 2.	
DR	PROSITE; PS00063; ALDO-KETO_REDUCTASE_3; UNKNOWN_1.	
KW	ATP-binding.	
SK	SEQUENCE 1532 AA; 171557 MW; C18F3554DD85732A CRC64;	
QY	Query Match	70.0%; Score 6932.5; DB 2; Length 1532;
DB	Best Local Similarity	87.3%; Pred. No. 0;
DB	Matches 1338; Conservative 100; Mismatches 92; Indels 3; Gaps 3;	
QY	1	MALRFGCSADGSDPLNDNNVTNTSPDFTKCFQNTLVWVPCFVLWACPFYFLYLSRH 60
DB	1	MALRFGCSDDGSDPLNDNNVTNTSPDFTKCFQNTLVWVPCFVLWACPFYFLYLSRH 60
QY	61	DRGYQTMTPLNKTALGFLMWICWADLFYSFWRSGRGIPLAPVFLVSPILLGITLLA 120
DB	61	DRGYQMTLHNKAKTALGFLMWICWADLFYSFWRSGRGIPLAPVFLVSPILLGITLLA 120
QY	121	TFLIQLRRKGVSIGIMTLFWLVALVCALAILRSKIMTALKBAQVDFLDFITFYVFS 180
DB	121	TFLIQFERRRKGVSIGIMTLFWLVALVCALAILRSKIMTALKBAQVDFLDFITFYVFS 180
QY	181	LLLQLVLSCFSDSPFLSETHDPNCPESASFLSRTITFWITGLIVRGYRQPLEGSD 240
DB	181	LVFQLVLSCFSDSPFLSETHDPNCPESASFLSRTITFWITGLIVRGYRQPLEGSD 240
QY	241	LWSLNKEDTSQVVPVLVKNWKKCAKTRKQPVKVVS - SKDPAQPKSSKVDANEVEA 299
DB	241	LWSLNKEDTSQVVPVLVKNWKKCAKTRKQPVKVVS - SKDPAQPKSSKVDANEVEA 299
QY	300	LIVKSPQKEMNPSLFKVLKTFPGVFLMSFFKAIHDMFSGPOIKLLIKFVNDTKAP 359
DB	301	LIVKSSKDRDPSLFKVLKTFPGVFLMSFFKAIHDMFSGPOIKLLIKFVNDTKAP 360
QY	360	DWQGYFTVLLFVTAQLTLVHGVPHICFVSGMRITKAVTGAVYRKALVITNSARKSST 419
DB	361	DWQGYFTVLLFVTAQLTLVHGVPHICFVSGMRITKAVTGAVYRKALVITNSARKSST 420
QY	420	VGEIVNLSVDAQRFMDLATYINIMWSAPLOVILALLVLLNLGSPVLAVGVMLVMPV 479
DB	421	VGEIVNLSVDAQRFMDLATYINIMWSAPLOVILALLVLLNLGSPVLAVGVMLVMPV 480
QY	480	NAVAMTKTKTQVAHMKSKONRIKLMNEILNGIKVLKLYAWELAFKQKVLAIROEELKVL 539
DB	481	NAVAMTKTKTQVAHMKSKONRIKLMNEILNGIKVLKLYAWELAFKQKVLAIROEELKVL 540
QY	540	KKSAYLSAVGTFTWVCTPFLVALCTFAVYVTTIDNNILDACTAFVSLALFNILRFPNLL 599
DB	541	KKSAYLAAGVGTFTWVCTPFLVALCTFAVYVTTIDNNILDACTAFVSLALFNILRFPNLL 600
QY	600	PWVISSIVQASVSLKRLIFLSHEELEPDSIERPVPKDGCGTNSITVKNATFTWARSDDP 659
DB	601	PWVISSIVQASVSLKRLIFLSHEELEPDSIERPVPKDGCGTNSITVKNATFTWARSDDP 660
QY	660	TLNGITFSPICALVAVVGVQVCGCKSSLLSALLAEMDKVEGHVAKGSAVVPQAWLQN 719
DB	661	TLNGITFSPICALVAVVGVQVCGCKSSLLSALLAEMDKVEGHVAKGSAVVPQAWLQN 720
QY	720	DSLRENILFGCOLPEPYRSVIOACALLPDLLEILPSSGDRTEIGEGVNLGGQKQVSLA 779
DB	721	DSLRENILFGCOLPEPYRSVIOACALLPDLLEILPSSGDRTEIGEGVNLGGQKQVSLA 780
QY	780	RAVYSNADILYLFDDPLSAVDAGKHIFENVIKPKMLKNKTRILVTHSMYILPQVDVII 839
DB	781	RAVYNSDIYLLDDPLSAVDAGKHIFENVIKPKMLKNKTRILVTHSMYILPQVDVII 840
QY	840	VMSGKISEMGSYQELLARDGAFALRTVASTEQEOBAENGVTGVSGPKAKOMENG 899
DB	841	VMSGKISEMGSYQELLARDGAFALRTVASTEQEOBAENGVTGVSGPKAKOMENG 900
QY	900	MLVTSAGKQLQROLSSSSSYSGDISRHNSSTABEQKAEKKEBTWKLMEADKAQTGVK 959
DB	901	ILVTDVAVGKPLQRHLSNSSSHSVTNQHGSTABLQKS - GVKEETWKLMEADKAQTGVK 959
QY	960	LSVYWDYMKATGLFISFLFMCNHSVLSALSNYMLSLWTD - PIVNGTQEHTKVRLSV 1018
DB	960	LSVYWNKALGLCISFLSFLFMCNHSVLSALSNYMLSLWTD - PIVNGTQEHTKVRLSV 1019
QY	1019	YGALGISQIAVFGYSMAVSGIGILASRCLHVDLHLSILRSPMSFFERTSGNLVNRFSK 1078
DB	1020	YGALGILQGVAVFGYSMAVSGIGIFASRLHLDLQNLVLRSPMSFFERTSGNLVNRFSK 1079
QY	1079	ELDTVDSMIPEVIMKMGSLFNVIGACIVILLATPIAAIIIPPLGLYFFVQRYVASSR 1138
DB	1080	ELDTVDSMIPEVIMKMGSLFNVIGAVIIILLATPIAAVIPPPLGLYFFVQRYVASSR 1139
QY	1139	QKRLSVSRSPVYSHENETLLGVSVIRAFEEQERFIHQSDLKVDENOKAYYPSIVANRW 1198
DB	1140	QKRLSVSRSPVYSHENETLLGVSVIRAFEEQERFIHQSDLKVDENOKAYYPSIVANRW 1199
QY	1199	LAVLECVGNCIVFAALFAVIRSHLSAGLVLSYSLSQVTTYTLNWLVRMSSEMETNI 1258
DB	1200	LAVLECVGNCIVFAALFAVIRSHLSAGLVLSYSLSQVTTYTLNWLVRMSSEMETNI 1259
QY	1259	VAVERLKEYSETEKAPWQIQETAPSPSWQVGRVFRNCLRYREDLDFVLRHINTIN 1318
DB	1260	VAVERLKEYSETEKASWQIQETAPSPSWQVGRVFRNCLRYREDLDFVLRHINTIE 1319
QY	1319	GGEKVGIVGRTGAGKSLTLGLFRINESAGEIIDIIGINIAKIGLHDLRFKTIIPDPV 1378
DB	1320	GGEKVGIVGRTGAGKSLTLGLFRINESAGEIIDIIGINIAKIGLHDLRFKTIIPDPV 1379
QY	1379	LFSGSLRMNLDPPFSQVSDSEVWTSLELAHLKDFVSALPKLDHCEAGGENLSVGORQV 1438
DB	1380	LFPSGLRMNLDPPFSQVSDSEVWTSLELAHLKDFVSALPKLDHCEAGGENLSVGORQV 1439
QY	1439	CLARALLRKTKILVLDENATVLDLTDLLIOSTIRTOFQEDCTVLTIAHRLNTIMDYTRVI 1498
DB	1440	CLARALLRKTKILVLDENATVLDLTDLLIOSTIRTOFQEDCTVLTIAHRLNTIMDYTRVI 1499
QY	1499	VLDKGEIQEYCAPSDLIQQRGLFYSSMAKADAGLV 1531
DB	1500	VLDKGEIQEYCAPSDLIQQRGLFYSSMAKADAGLV 1532
QY	RESULT 13	
DB	Q810G9	
ID	Q810G9	PRELIMINARY; PRT; 1523 AA.
AC	Q810G9	
DT	01-JUN-2003	(TREMBlrel. 24, Created)
DT	01-JUN-2003	(TREMBlrel. 24, Last sequence update)
DT	01-MAR-2004	(TREMBlrel. 26, Last annotation update)

DE ATP-binding cassette protein Cl variant A.  
GN Name=Abcc1;  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Sprague-Dawley; TISSUE=Spleen;  
RA Takayanagi S., Iehikawa T.;  
RL Submitted (NOV-2002) to the EMBL/GenBank/DDAJ databases.  
CC -!- SIMILARITY: Belongs to the ABC transporter family.  
DR EMBL; AY174892; AA044983.1; -;  
DR HSSP; P08716; IMT0.  
DR GO; GO:0016021; C: integral to membrane; IEA.  
DR GO; GO:0005524; F: ATP binding; IEA.  
DR GO; GO:0042626; F: ATPase activity, coupled to transmembrane m. . .; IEA.  
DR GO; GO:0000166; F: nucleotide binding; IEA.  
DR GO; GO:0005215; F: transporter activity; IEA.  
DR GO; GO:0006810; P: transport; IEA.  
DR InterPro; IPR003593; AAA ATPase.  
DR InterPro; IPR011527; ABC membrane 1.  
DR InterPro; IPR001140; ABC\_TM\_transp.  
DR InterPro; IPR003439; ABC\_transp.  
DR InterPro; IPR001395; ABC\_transp.  
DR InterPro; IPR005292; MRP\_assoc.  
DR Pfam; PF00664; ABC membrane; 2.  
DR Pfam; PF00005; ABC tran; 2.  
DR ProDom; PD000006; ABC transporter; 2.  
DR SMART; SM00382; AAA; 2.  
DR TIGRFAMs; TIGR00957; MRP\_assoc\_pro; 1.  
DR PROSITE; PS0929; ABC TMIF; 2.  
DR PROSITE; PS0211; ABC\_TRANSPORTER\_1; 2.  
DR PROSITE; PS0893; ABC\_TRANSPORTER\_2; 2.  
DR PROSITE; PS00663; ALDO\_KETO\_REDUCTASE\_3; UNKNOWN\_1.  
KW ATP-binding.  
SQ SEQUENCE 1523 AA; 170503 MW; B40337051A1CB9C6 CRC64;

Query Match 69.6%; Score 6892; DB 2; Length 1523;  
Best Local Similarity 87.0%; Pred. No. 0;  
Matches 1333; Conservative 98; Mismatches 90; Indels 12; Gaps 4;

QY 1 MALRGCSDGSDPLDNDVNTWNTNPDFTKCFQNTLVWPCFYLMACFPFYLYLSRH 60  
DB 1 MALRGCSDGSDPLDNDVNTWNTNPDFTKCFQNTLVWPCFYLMACFPFYLYLSRH 60

QY 61 DRGYIQMTPLNKTALGFLLMIVCWADLFYFWERSRGIPLAPVLYSPTLIGITLLA 120  
DB 61 DRGYIQMTPLNKTALGFLLMIVCWADLFYFWERSRQGLLAPVLYSPTLIGITLLA 120

QY 121 TPLIQLERRKGVQSSGIMLTFWLVLCALALRSKIMTALKEDAQVDLFRDITFVYVS 180  
DB 121 TPLIQLERRKGVQSSGIMLTFWLVLCALALRSKIMTALKEDAQVDLFRDITFVYVS 180

QY 181 LLLIQLVLCFSDRSLPSETIHPNCPPESSASFLSRITFWITGLIVRGYRQPLEGSD 240  
DB 181 LVFIQLVLCFSDRSLPSETIHPNCPPESSASFLSRITFWITGLIVRGYRQPLEGSD 240

QY 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPVVYS-SKDPAPQKSSKVDANEEVEA 299  
DB 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPVVYS-SKDPAPQKSSKVDANEEVEA 300

QY 300 LTVKSPQKWNPSLFVLYKTFPGPYFLMSFFPKALHDLMMFSGPOILLIKFVNDTRAP 359  
DB 301 LTVKSHKDRDPSLFVLYKTFPGPYFLMSFLYKALHDLMMFAGPILLELIINFVNDREP 360

QY 360 DWQGYFTVLLFVTLACQTLVHQYFHCIFVSGMRKTAIVGAVYRKALVITNSARKSST 419  
DB 361 DWQGYLYTALLFVSAQTLVHQYFHCIFVSGMRKTAIVGAVYRKALVITNSARKSST 420

QY 420 VGEIVNLMSVDAQRFMDLATYINMIWSAPLQVILALYLWNLGSPVLGAVVWLVMPV 479  
DB 421 VGEIVNLMSVDAQRFMDLATYINMIWSAPLQVTLALYFLWNLGSPVLGAVVWLVMPV 480

QY 480 NAVMAKTKTYQVAHMKSKONRIKLMNEILNGIKVLKLYAWELAFKDKVLAIROBELKVL 539  
DB 481 NAVMAKTKTYQVAHMKSKONRIKLMNEILNGIKVLKLYAWELAFKDKVMNIRQBELKVL 540

QY 540 KKSAYLSAVGTFVWCTPFLVALCTFAVYVVTIDENNILDQAQTAFAVSLALFNILRPLNLT 599  
DB 541 KKSAYLAAGVTFVWCTPFLVALSTFAVFTVDEKNILDAKKAFVSLALFNILRPLNLT 600

QY 600 PMWISSIVQASVSLKRLRIFLSHBELEPDSIERPVDKGGGTSITVRNATFTWARSDDP 659  
DB 601 PMWISSIVQASVSLKRLRIFLSHBELEPDSIERWSIKDGGGWNSTVKNATFTWARDDEP 660

QY 660 TLNGITFSIPGALVAVVGVQCGKSSLLSALLAEMDKVEGHVAIKGSVAVVYVQQAQWION 719  
DB 661 TLNGITFAIPDGLVAVVGVQCGKSSLLSALLAEMDKVEGHVTLKGSVAVVYVQQAQWION 720

QY 720 DSLRENILFGQLEPEYRSVIOACALPDLEILPSGDRTEIGEGKGNLSGQKORVSLA 779  
DB 721 DSLRENILFGRLPQEHYKAVMEACALLPDLEILPSGDLTEIGEGKGNLSGQKORVSLA 780

QY 780 RAVYSNADIYLFDDPLSAVDHVGHHFENVIGPKGMLKNKTRILVTHSMSYLPQVDVII 839  
DB 781 RAVYNSDIYLLDPLSAVDHVGHHFENVIGPKGMLKNKTRILVTHGSIYLPQVDVII 840

QY 840 VMSGKISEMGSYQELLARDGAFLEFRTYASTQEOBAENGVTGVSQPGKEAKOMENG 899  
DB 841 VMSGKISEMGSYQELLARDGAFLEFRTYASTQEOBAENGVTGVSQPGKEAKOMENG 900

QY 900 MLVTDGAKLQRLSSSSSYSGDISRHNSHTAELOKAEKEETWKLMEADKAOTGVK 959  
DB 901 ILVTDGAKPL-----HSVVTNQHSSTAELOKS-GVKETWKLMEADKAOTGVK 950

QY 960 LSVYDYMKAIGLFTISFLSIFLFCMCHVSNALSNYLSLWTDG-PVANGTQSHTKVRLSV 1018  
DB 951 LSVYDYMKAIGLFTISFLSIFLFCMCHVSNALSNYLSLWTDGPAVNGTQENRFRSLV 1010

QY 1019 YGALGISQGIADVFGYSMAVSIIGGILASRCLHVDLLHSILRSPMSFFERTPGNLRNFSK 1078  
DB 1011 YGALGILQGVAVFGYSMAVSIIGGIFASRRLHLDLQNLRSFMSFFERTPGNLRNFSK 1070

QY 1079 ELDTVDSMIPVINKFMGSLFNIVGACIVILLATPIAIIIPPLGLIYFFVQRFVYASSR 1138  
DB 1071 ELDTVDSMIPVINKFMGSLFSVIGAVIIILATPIAIIIPPLGLVYFFVQRFVYASSR 1130

QY 1139 QKRLSVSRSPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENKAYYPSIVANRW 1198  
DB 1131 QKRLSVSRSPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENKAYYPSIVANRW 1190

QY 1199 LAVRLECYGNCIYVLAALFAVIRSHLSAGLVGLSVSYSLQITAYLNLVLRMSSEMETNI 1258  
DB 1191 LAVRLECYGNCIYVLAALFAVIRSHLSAGLVGLSVSYSLQITAYLNLVLRMSSEMETNI 1250

QY 1259 VAVERLKEYSETEKAPWQIOETAPPSWPQVGRVFRNYCLRYREDLDFVLRLHNVITIN 1318  
DB 1251 VAVERLKEYSETEKASWQIOETAPPSWPHSGRVEFRDYCLRYREDLDFVLRLHNVITIE 1310

QY 1319 GGEKGVIGRTGAGKSSLTGLFRINESAEBEIIIDGINIAKIGLHDLRFKTIIPQDPRV 1378  
DB 1311 GGEKGVIGRTGAGKSSLTGLFRINESAEBEIIIDGINIAKIGLHDLRFKTIIPQDPRV 1370

QY 1379 LFGSLRNLDLPPFSQYSDDEEYVMTSLAHLKDFVSALPDKLDHECAEGENLSVGQRQLV 1438  
DB 1371 LFGSLRNLDLPPFSQYSDDEEYVMTSLAHLKDFVSALPDKLDHECAEGENLSVGQRQLV 1430

QY 1439 CLARALLKTKILVLDEATAVADLETDDLIQSTIRTPQEDCTVLIHRLNTIMDYTRVI 1498  
DB 1431 CLARALLKTKILVLDEATAVADLETDDLIQSTIRTPQEDSTVLTIAHRLNTIMDYTRVI 1490

QY 1499 VLDKGEI0EYGAAPSLLQORGLFYSMKADAGLV 1531  
DB 1491 VLDKGEI0EYGAAPSLLQORGLFYSMKADAGLV 1523

```

RESULT 14
Q68CP7 PRELIMINARY; PRT; 1215 AA.
ID Q68CP7;
AC Q68CP7;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein DKFZp781G125.
GN Name=DKFZp781G125;
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_
RP SEQUENCE FROM N.A.
RC TISSUE=Amygdala;
RG The German cDNA Consortium;
RA Koehrer K., Beyer A., Mewes H.W., Weil B., Amid C., Oeanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: CR749835, CAH10691.1;
DR InterPro: IPR003593; AAA ATPase.
DR InterPro: IPR001140; ABC TM transport.
DR InterPro: IPR003439; ABC transporter.
DR InterPro: IPR005292; MRP assoc.
DR InterPro: IPR000719; Prot kinase.
DR Pfam: PF00564; ABC membrane; 2.
DR Pfam: PF00005; ABC_tran; 2.
DR ProDom: PD000006; ABC transporter; 2.
DR SMART: SM00382; AAA; 2.
DR TIGRFAMs: TIGR00957; MRP assoc pro; 1.
DR PROSITE: PS00211; ABC TRANSPORTER 1; 2.
DR PROSITE: PS00893; ABC TRANSPORTER 2; 2.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
KW ATP-binding; Hypothetical protein.
SQ SEQUENCE 1215 AA; 134926 MW; E490AD1F5268F72E CRC64;

Query Match 61.3%; Score 6075; DB 2; Length 1215;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1202; Conservative 0; Mismatches 3; Indels 10; Gaps 1;

QY 327 MSFFPKAIHDLMMFSGPQILKLLIKFVNDTKAPDWQGYFYTVLLFVTACLOTLLVHOYFH 386
Db 1 MSFFPKAIHDLMMFSGPQILKLLIKFVNDTKAPDWQGYFYTVLLFVTACLOTLLVHOYFH 60

QY 387 ICFVSGMERIKTAVIGAVYRKALVITNSARKSTVGEIYNLMSVDAQRFMDLATYINMWS 446
Db 61 ICFVSGMERIKTAVIGAVYRKALVITNSARKSTVGEIYNLMSVDAQRFMDLATYINMWS 120

QY 447 APLQVILALYLMLNLGPSVLGAVVMVLPVNAVMAKTKTYQVAHMKSKDNRIKLMN 506
Db 121 APLQVILALYLMLNLGPSVLGAVVMVLPVNAVMAKTKTYQVAHMKSKDNRIKLMN 180

QY 507 EILNGIKVLKYAMELAFKDLAIROBELKVLKKSAYLSAVGTFTWCTPFLVALCTFA 566
Db 181 EILNGIKVLKYAMELAFKDLAIROBELKVLKKSAYLSAVGTFTWCTPFLVALCTFA 240

QY 567 VVVTIDENNIIDAQTAFLSLALENLREPTNAILPMTISSIYQASVSLKRLIFLSHEELE 626
Db 241 VVVTIDENNIIDAQTAFLSLALENLREPTNAILPMTISSIYQASVSLKRLIFLSHEELE 300

QY 627 PDSIERRPVKDGGTNSITVRNATFTWARSDDPTINGITFTSIPGALVAVVGQVCGCKSS 686
Db 301 PDSIERRPVKDGGTNSITVRNATFTWARSDDPTINGITFTSIPGALVAVVGQVCGCKSS 360

QY 687 LLSALLAEMDKVEGHVAIKGSVAVYPOQAWTQNDLSRENILFGQLSEPYRYSVIQACAL 746
Db 361 LLSALLAEMDKVEGHVAIKGSVAVYPOQAWTQNDLSRENILFGQLSEPYRYSVIQACAL 420

QY 747 LPDLIELPSGDRTEIGEKGVLNSGQKORVSLARAVYSNADIYLFDDPLSAVDHVGKHI 806
Db 421 LPDLIELPSGDRTEIGEKGVLNSGQKORVSLARAVYSNADIYLFDDPLSAVDHVGKHI 480

QY 807 FENVIGPKMLKNKTRILVTHSMSTYLPQDVIIIVMSGGKISMGSGYQELLARDGAFBFL 866
Db 481 FENVIGPKMLKNKTRILVTHSMSTYLPQDVIIIVMSGGKISMGSGYQELLARDGAFBFL 540

QY 867 RTYASTEEOQDAEEN-----GVTVSGPGKEAKOMENGLMTDSAGKQLQRQLSS 916
Db 541 RTYASTEEOQDAEENGVTVSGPGKEAKOMENGLMTDSAGKQLQRQLSS 600

QY 917 SSSYSGDISRHHNSTAEQKAEAKKEETWKLMEADKAQTGVKLSVYWDYMYKAIGLFTSF 976
Db 601 SSSYSGDISRHHNSTAEQKAEAKKEETWKLMEADKAQTGVKLSVYWDYMYKAIGLFTSF 660

QY 977 LSIPLFMCNHNVSALASNYWLSLWTDPIVNGTOEHTKVRLSVYGALGISQGIAPVGYGMA 1036
Db 661 LSIPLFMCNHNVSALASNYWLSLWTDPIVNGTOEHTKVRLSVYGALGISQGIAPVGYGMA 720

QY 1037 VSIGGILASRCLHVDLLHLSILRSPMSFPPTSGNLVNRFSKELDTVDSDMIPVTKMFWG 1096
Db 721 VSIGGILASRCLHVDLLHLSILRSPMSFPPTSGNLVNRFSKELDTVDSDMIPVTKMFWG 780

QY 1097 SLFNVIGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQKRLSVSRSPVYSHFN 1156
Db 781 SLFNVIGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQKRLSVSRSPVYSHFN 840

QY 1157 ETLIGVSVIRAFERQERFIHQSDLKVDENQKAYYPSIVANRWLAVRLECVGNCIVLFAAL 1216
Db 841 ETLIGVSVIRAFERQERFIHQSDLKVDENQKAYYPSIVANRWLAVRLECVGNCIVLFAAL 900

QY 1217 FAVISRHSLSAGLVGLSVSLQVTVYTNLVRMSSEMETNIVAVERKEYSETEKAPW 1276
Db 901 FAVISRHSLSAGLVGLSVSLQVTVYTNLVRMSSEMETNIVAVERKEYSETEKAPW 960

QY 1277 QIQTAPPSSWPQVGRVEFRNYCLRYREDLDFVLRLHNTINGGKVGIVGRTGAGKSSL 1336
Db 961 QIQTAPPSSWPQVGRVEFRNYCLRYREDLDFVLRLHNTINGGKVGIVGRTGAGKSSL 1020

QY 1337 TLGLFRINESAGEIIIDGINIAKIGHDLRFKTIIPQDPVLFSGSLRNWLDPPSOYSD 1396
Db 1021 TLGLFRINESAGEIIIDGINIAKIGHDLRFKTIIPQDPVLFSGSLRNWLDPPSOYSD 1080

QY 1397 EEVWTSLELAHLKDFVSALPKDHECAEGENLSVGQRLVCLARALLRKTILVLDEA 1456
Db 1081 EEVWTSLELAHLKDFVSALPKDHECAEGENLSVGQRLVCLARALLRKTILVLDEA 1140

QY 1457 TAAVDLETDDLIOSTIRTFQEDCTVLTIAHRLNTIMDYTRVIVLDKGEIOEYGAPSDLLQ 1516
Db 1141 TAAVDLETDDLIOSTIRTFQEDCTVLTIAHRLNTIMDYTRVIVLDKGEIOEYGAPSDLLQ 1200

QY 1517 QRGLFYSMAXDAGLV 1531
Db 1201 QRGLFYSMAXDAGLV 1215

RESULT 15
Q80ZK8 PRELIMINARY; PRT; 1519 AA.
ID Q80ZK8;
AC Q80ZK8;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Similar to ATP-binding cassette, sub-family C (CFTR/MRP), member 3 (Fragment).
DE (Fragment).
GN Name=Abcc3;
OS Mus musculus (Mouse);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Liver;
RA Strausberg R.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.

```

-!- SIMILARITY: Belongs to the ABC transporter family.

CC EMBL; BC048825; AAH48825.1; -.  
 DR HSP; P26361; 1R0Z.  
 DR GO; MGI:1923658; Abcc3.  
 DR GO; GO:0005887; C-integral to plasma membrane; IDA.  
 DR InterPro; IPR003593; AAA ATPase.  
 DR InterPro; IPR011527; ABC transmembrane 1.  
 DR InterPro; IPR001140; ABC TM transp.  
 DR InterPro; IPR003439; ABC transporter.  
 DR InterPro; IPR005292; MRP\_assoc.  
 DR Pfam; PF00664; ABC membrane; 2.  
 DR Pfam; PF00005; ABC\_tran; 2.  
 DR ProDom; PD00006; ABC transporter; 2.  
 DR SMART; SM00382; AAA; 2.  
 DR TIGRFAMs; TIGR00957; MRP assoc pro; 1.  
 DR PROSITE; PS00929; ABC TMIF; 2.  
 DR PROSITE; PS00211; ABC TRANSPORTER 1; 2.  
 DR PROSITE; PS00893; ABC\_TRANSPORTER\_2; 2.  
 KW ATP-binding.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 1519 AA; 168648 MW; 72A5151A862EC948 CRC64;  
 Query Match 45.7%; Score 4522; DB 2; Length 1519;  
 Best Local Similarity 56.2%; Pred.No. 2.5e-270;  
 Matches 865; Conservative 269; Mismatches 366; Indels 38; Gaps 8;  
 QY 8 SADGSDPLDWNVTWNTGNPDFTKCFQNTVLVWVPCFLWACFPFFLYLGRHGRGIOM 67  
 DB 4 SGELGSKFWDNLSITTYTPDLTPCFQNSLLAWFCIVLWALPCYFLRHHQGIYVL 63  
 QY 68 TPLNKTATGALLVIVCWADLFYFWSRSGIFLAPVFLVPTLLGTTLLATLFIQLE 127  
 DB 64 SWLSRLKLTALGVLLWCVSWDLFYFHLGHGSSPAPVFFVPLVVGITMLLATLLIQYE 123  
 QY 128 RRGVQSSGIMLTFMLVALCALILRSKIMTALKEDAQVDFRDIITYVFSLLIIQIV 187  
 DB 124 RLRGVQSSGVLLIFWLLCAIIPFRSKILSALAEGKILDPFRFTTYIYPALVFCALI 183  
 QY 188 LSCFSDRPLFSETHDNPCESSASFLSRITFWITGLIVRGYRQLEGSDLWSLNKE 247  
 DB 184 LSCFKEKPLSPENLDINPCPEASAGPFRSLFWFTRFALIGYRRPLEDBDLWSLSE 243  
 QY 248 DTSEQVPLVKNWKECAKTRKQPVVYVSKDPAQPKESKVDANEVEBALIVKSPQK 307  
 DB 244 DCSHRVQRLLEAMQKQ---QNASGQOTATAEPKIPGE-----DAVLLKPRPK 289  
 QY 308 EWNPSLFVLYKTPGYPFLMSPPFKAIHDLMMFSGPOLIKLLIKFVNDTKAPDWQGYFT 367  
 DB 290 SKQSPSLRALVTRFTTSSLLMSACFNLIQNLGFPNPQLLSILIRFISDPTPTWGFLLA 349  
 QY 368 VLLFVYACQLVLAHQYFHCIFVSGMRITKTAIVGAYRKALVITNSARKSVTVGEIVNLM 427  
 DB 350 GLMFLSSMTQTLILHQYHCFVWALRLRTALIGVYRKALVITNSVKRESTVGEVNLN 409  
 QY 428 SVDAQRFMDLATYINMINSAPQVILALYLWNLNGLPSVLAGVAVMLVMPVNAVMAKT 487  
 DB 410 SVDAQRFMDVSPFINLLSAPQVILALYLWNLNGLPSVLAGVAVIVLLIPLNGAVSKM 469  
 QY 488 KTYQVAHMSKDNRIKLMNEILNGIKVLYAWELAPKQVLAIRQBELKVLKSAIYLSA 547  
 DB 470 KTYQVKQKFKDSRIKLMSEILNGIKVLYAWEPSPLEQVKGIRQSELQQLRKGALQA 529  
 QY 548 VCTFTWCTPPLVALCTFAVYVYTIENNILLDAQTAFLVSLALFNILRFPNLLPMVISSIV 607  
 DB 530 ISTFTWICTPPLVLTITLVGVYVYVDESNDVDEKAFVSLNFIUKIPLNMLPQLISGLT 589  
 QY 608 QASVSLKRLRIFLSHEEPEPDSIERPRVKDGGTNSITVRNATFTWASDPPTLNGITFS 667  
 DB 590 QASVSLKRIQDFLNQNELDPQCVERKTISPQ---YAIITHNGTFTWAQDLRPTLHSLNIQ 646  
 QY 668 IPEGALVAVGVQCGKSGLLSALLAEMDKVEGHVAIKGSVAVYVQQAQWQNDLSRENIL 727  
 DB 647 IPKGAIVAVGVQCGKSLVALLGEMEKLEGVVSVKGSVAVYVQQAQWQNDLSRENIL 706

QY 728 FGCOLEEPYRSVIAQACALLPDLPEILPSGDRTEIGEKGWNLGGQKQVRSARAVYSNAD 787  
 DB 707 FGQPMNPKRYQOALETALLADLVLPGGDQTEIGEKGINLGGQKQVRSARAVYSNAD 766  
 QY 788 IYLDPPDPLSAVDHVGKHFENVIGPKMLKXKTRILYTHSMYLPQVDVIVVMSGGKIS 847  
 DB 767 IFLLDDPLSAVDHSHVAKHIFDQVIGPEGLVAGTRVLVTHGISFLPQDFDIIVLAGGQVS 826  
 QY 848 EMGSYQELLARDGAPAEPLRYASTEQDQAEENGVTGVSQPKAKOMENGML-----VT 903  
 DB 827 EMGHYSALLQHDGSAFANPLRYAPDEQDEH-----ALQANAEVLLLEDLTSLTHDILT 881  
 QY 904 DS-----AGQLOLQOLSSSSSSSGDIS-----RHNSTABLOKAEKKEETWKLMEADK 952  
 DB 882 DNEPAIVEVRKQFMREMSLSS-EGEVQNRTPMKKHTNSLEKEALVTTKETGALIKEBI 940  
 QY 953 AQTQVKLSVVDYWKALIGLIFSLFPLFCNHNVSALASNYWLSLWTDVINGTQBEHT 1012  
 DB 941 AETGNVLSVVDYAKSMGLCTTILSICLLYGGQSAANAIGANVWLSAMSNDABEHQOQNK 1000  
 QY 1013 KVLRSVYGALGISQGIADVFGYSMAVSIIGILASRCLHVDLLHLSILRSPMSFFERTPSGNL 1072  
 DB 1001 SVRLGVYAALGILQGLLWMLSAFTWVGAIQARLLHEALLHNKIRSPQSFDTTPSGRI 1060  
 QY 1073 VNRFSKELDTYDSMIPEVIRKMFMSLNFVIGACIVILLATPIAAIIIPPLGLIYFVQRF 1132  
 DB 1061 LNRFESKIDYIDEVLAPTILMLNSFFTSISTIMVIVASTPLFMVWVPLAVLYGFVQRF 1120  
 QY 1133 YVASRQLKLESVRSRSPVYSHFNETHLLGVSVIRAFEEQERFIHQSDKLVNENQAYYPS 1192  
 DB 1121 YVATSRQLKLESISRSPISFHSFSETVTGTSVIRAYGRIOQDFKVLSDTKVDNNQKSSYP 1180  
 QY 1193 IVANRWLAVERLKEVSETEKEAPQIOETAPPSSWPQVGRVFRNYCLRYREDLDFVLRH 1252  
 DB 1181 IASNRWLGVRHVEFGNVCVLLFAALFAVIGRNSLAPGLVGLSVYALQVTMALNWMIRMS 1240  
 QY 1253 EMETNIVAVERLKEVSETEKEAPQIOETAPPSSWPQVGRVFRNYCLRYREDLDFVLRH 1312  
 DB 1241 DLESNIIAVERVKEYSKTKTEAPVWVESNRAPEGWPTGRGWVFRNYSVRYRPGLELVKN 1300  
 QY 1313 INVINGEKVIGVGRGTGAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDIRPKITI 1372  
 DB 1301 VTVHVGGEKVGIVGRGTGAGKSSMTCLFRILEAAGEIVIDGLVAHIGLHDLRSQITI 1360  
 QY 1373 IPQDPVLPFSGSLRWNLDPFQVSDVEVWTSLELAHLKDFVSALPKLDHECAEGGENLSV 1432  
 DB 1361 IPQDPVLPFSGSLRWNLDPFQVSDVEVWTSLELAHLKDFVSALPKLDHECAEGGENLSV 1420  
 QY 1433 GQRLVCLARALLRKTKILVLDATAAVIDLETDLLIQSTIRTPQEDCTVLTIAHRLNTIM 1492  
 DB 1421 GQRLVCLARALLRKTKILVLDATAAVIDLETDLLIQSTIRTPQEDCTVLTIAHRLNTIM 1480  
 QY 1493 DYTRVIVLDKGEIOEYGAAPSDLLQORGLFYMAKDAGL 1530  
 DB 1481 DYNRVVLDKGVAAEFDSFVNLIAGGIFYGMAKDAGL 1518

Search completed: March 18, 2005, 11:12:16  
 Job time : 212.942 secs

**THIS PAGE BLANK (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2005, 10:57:20 ; Search time 40.7091 Seconds  
(without alignments)  
3570.253 Million cell updates/sec

Title: US-10-665-283-8

Perfect score: 10016

Sequence: 1 MALRGFCSDGSDPLMDWNV.....RSVAVAKPKFSISPDLS 1947

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/prodata/1/aaa/5A COMB.pep:\*\*
- 2: /cgn2\_6/prodata/1/aaa/5B COMB.pep:\*\*
- 3: /cgn2\_6/prodata/1/aaa/6A COMB.pep:\*\*
- 4: /cgn2\_6/prodata/1/aaa/6B COMB.pep:\*\*
- 5: /cgn2\_6/prodata/1/aaa/PCTUS COMB.pep:\*\*
- 6: /cgn2\_6/prodata/1/aaa/backfiles1.pep:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7860	78.5	1531	1	US-08-463-092B-4
2	7860	78.5	1531	2	US-08-462-109A-4
3	7860	78.5	1531	2	US-08-460-907B-4
4	7860	78.5	1531	3	US-08-463-179A-4
5	7860	78.5	1531	3	US-08-461-384B-4
6	7860	78.5	1531	4	US-09-647-140B-19
7	7849	78.4	1531	1	US-08-141-893-2
8	7849	78.4	1531	1	US-08-463-092B-2
9	7849	78.4	1531	2	US-08-462-109A-2
10	7849	78.4	1531	2	US-08-460-907B-2
11	7849	78.4	1531	3	US-08-463-179A-2
12	7849	78.4	1531	3	US-08-461-384B-2
13	7849	78.4	1531	3	US-08-407-207A-2
14	7002.5	69.9	1528	1	US-08-463-092B-6
15	7002.5	69.9	1528	2	US-08-462-109A-6
16	7002.5	69.9	1528	3	US-08-460-907B-6
17	7002.5	69.9	1528	3	US-08-463-179A-6
18	7002.5	69.9	1528	3	US-08-461-384B-6
19	4487.5	44.8	1527	4	US-09-647-140B-6
20	4487.5	44.8	1530	4	US-09-647-140B-33
21	3391.5	33.9	1503	4	US-09-647-140B-8
22	3385.5	33.8	1503	4	US-09-792-616-3
23	3318.5	32.1	1498	4	US-09-792-616-9
24	2374.5	23.7	1621	3	US-08-972-927-3
25	2349.5	23.5	1622	3	US-08-972-927-6
26	2266.5	22.6	1325	4	US-09-647-140B-2
27	2195.5	21.9	1261	3	US-09-439-313-538

28	2195.5	21.9	1261	4	US-09-636-215-538	Sequence 538, App
29	2195.5	21.9	1261	4	US-09-685-166A-538	Sequence 538, App
30	2195.5	21.9	1261	4	US-09-679-426-538	Sequence 538, App
31	2195.5	21.9	1261	4	US-09-759-143-538	Sequence 538, App
32	2195.5	21.9	1261	4	US-09-651-236-538	Sequence 538, App
33	2118	21.1	1581	3	US-08-726-320-3	Sequence 3, Appli
34	2118	21.1	1581	3	US-09-208-716-3	Sequence 3, Appli
35	2115	21.1	1228	3	US-09-439-313-537	Sequence 537, App
36	2115	21.1	1228	4	US-09-636-215-537	Sequence 537, App
37	2115	21.1	1228	4	US-09-685-166A-537	Sequence 537, App
38	2115	21.1	1228	4	US-09-679-426-537	Sequence 537, App
39	2115	21.1	1228	4	US-09-759-143-537	Sequence 537, App
40	2115	21.1	1228	4	US-09-651-236-537	Sequence 537, App
41	2082	20.8	1437	3	US-09-061-400-2	Sequence 2, Appli
42	2082	20.8	1453	3	US-09-001-273-2	Sequence 2, Appli
43	2082	20.8	1453	3	US-08-843-459A-2	Sequence 38, Appli
44	2081	20.8	1437	4	US-10-162-012-38	Sequence 4, Appli
45	2080	20.8	1437	4	US-09-647-140B-4	

ALIGNMENTS

RESULT 1  
US-08-463-092B-4  
; Sequence 4, Application US/08463092B  
; Patent No. 5766880  
; GENERAL INFORMATION:  
; APPLICANT: Cole, Susan P.C.  
; APPLICANT: Deeley, Roger G.  
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ENCODING  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS  
; STREET: Queen's University at Kingston  
; CITY: Kingston  
; STATE: Ontario  
; COUNTRY: CANADA  
; ZIP: K7L 3N6  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/463,092B  
; FILING DATE: 05-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/966,923  
; FILING DATE: 27-OCT-1992  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/029,340  
; FILING DATE: 8-MAR-1993  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/141,893  
; FILING DATE: 26-OCT-1993  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/407,207  
; FILING DATE: 20-MAR-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Steeg, Carol Miernicki  
; REGISTRATION NUMBER: 39,539  
; REFERENCE/DOCKET NUMBER: Q1546  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (613) 545-2342  
; TELEFAX: (613) 545-6853  
; INFORMATION FOR SEQ ID NO: 4:

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 1531 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-463-092B-4

Query Match      78.5%; Score 7860; DB 1; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGCSDGSDPLMDWNVNTWNTSNPDKFQNTLVWVPCFYLMACFPFYFLYLSRH 60
Db 1 MALRGCSDGSDPLMDWNVNTWNTSNPDKFQNTLVWVPCFYLMACFPFYFLYLSRH 60
QY 61 DRGYQMTPLNKTALGELLIVCWADLFYFWERSRGIFLAPVFLVSPDLLGITLLA 120
Db 61 DRGYQMTPLNKTALGELLIVCWADLFYFWERSRGIFLAPVFLVSPDLLGITLLA 120
QY 121 TFLIQLERRKGVOSSGIMLTFWLVALCALAIRSKIMTALKEDAQVDLFRDITFYVYFS 180
Db 121 TFLIQLERRKGVOSSGIMLTFWLVALCALAIRSKIMTALKEDAQVDLFRDITFYVYFS 180
QY 181 LLLIQLVSCFSDRSPLETHDNPCESSASFLSRITFWITGLIVRGYRQPLEGSD 240
Db 181 LLLIQLVSCFSDRSPLETHDNPCESSASFLSRITFWITGLIVRGYRQPLEGSD 240
QY 241 LWSLNKEDTSEQVPLVKNWKECAKTRKQPVVYVSKDPAQPKESKVDANEVEAL 300
Db 241 LWSLNKEDTSEQVPLVKNWKECAKTRKQPVVYVSKDPAQPKESKVDANEVEAL 300
QY 301 IVKSPQKWNPSLFVLYKTFGPFYLMGFFFKAIHDLMMFSGPQILKLLIKFVNDTKAPD 360
Db 301 IVKSPQKWNPSLFVLYKTFGPFYLMGFFFKAIHDLMMFSGPQILKLLIKFVNDTKAPD 360
QY 361 WQGYFYTVLLFVTAQLTLVHLQYPHICFVSGMRKTAIVGAVYRKALVITNSARKSSTV 420
Db 361 WQGYFYTVLLFVTAQLTLVHLQYPHICFVSGMRKTAIVGAVYRKALVITNSARKSSTV 420
QY 421 GEIVNLSMDAORFMDLATYINMIWSAPLOVTLALYLWLNIGPSVLGAVAVMLVMPVN 480
Db 421 GEIVNLSMDAORFMDLATYINMIWSAPLOVTLALYLWLNIGPSVLGAVAVMLVMPVN 480
QY 481 AVAMKTKTYQVAHMKSDNRKIKLMEILNGIKVLKYAWELAFKDKVLAIQBELKVLK 540
Db 481 AVAMKTKTYQVAHMKSDNRKIKLMEILNGIKVLKYAWELAFKDKVLAIQBELKVLK 540
QY 541 KSAYLSAVGTFTWCTPFLVALCTPAVYVTTDENNDILDAQTAFAVSLALFNILRPLNLP 600
Db 541 KSAYLSAVGTFTWCTPFLVALCTPAVYVTTDENNDILDAQTAFAVSLALFNILRPLNLP 600
QY 601 MVISSIVQASVSLKRLRIFLPSHEELEPDSIERRPVKDGGTNSITVRNATFTWARSDDPT 660
Db 601 MVISSIVQASVSLKRLRIFLPSHEELEPDSIERRPVKDGGTNSITVRNATFTWARSDDPT 660
QY 661 LNTGIFTSIPEGALVAVVQVCGCKSSLLSALLAEMDKVEGHVAIKGSVAYVPPQAWIQND 720
Db 661 LNTGIFTSIPEGALVAVVQVCGCKSSLLSALLAEMDKVEGHVAIKGSVAYVPPQAWIQND 720
QY 721 SIRENILFGCOLLEPYRSVTOACALLPDLILPSGDRTEIGKGVNLSGGQKORVSLAR 780
Db 721 SIRENILFGCOLLEPYRSVTOACALLPDLILPSGDRTEIGKGVNLSGGQKORVSLAR 780
QY 781 AVYSNADILYFDDPLSADVAHVGHIFENVIKPGMLKNKTRILVTHSMSYLPQVDVLIIV 840
Db 781 AVYSNADILYFDDPLSADVAHVGHIFENVIKPGMLKNKTRILVTHSMSYLPQVDVLIIV 840
QY 841 MSGGKISEMGVQELLARDAFAFLRYASTEQDAEENGVTGSGPGKEAKOMENGM 900
Db 841 MSGGKISEMGVQELLARDAFAFLRYASTEQDAEENGVTGSGPGKEAKOMENGM 900
QY 901 LVTSAGKOLQRLSSSSSSSGDISRRHNSAEQLQKAEKKEETWKLMEADKATQGVKL 960

```

```

RESULT 2
US-08-462-109A-4
; Sequence 4, Application US/08462109A
; Patent No. 5882875
; GENERAL INFORMATION:
; APPLICANT: Cole, Susan P.C.
; APPLICANT: Deeley, Roger G.
; TITLE OF INVENTION: METHODS FOR IDENTIFYING
; TITLE OF INVENTION: MULTIDRUG RESISTANT TUMOR CELLS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/462,109A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/966,923
; FILING DATE: 27-OCT-1992
; APPLICATION NUMBER: 08/029,340

```

Db	721	SURENLFGQLEEPYRSVIOACALLPDLLEILP	SGDRTEIGBKGNLSGGQKQRVSLR	780
Qy	781	AVYSNADIYLFDDPLSAVDHVGKHI	FENVIGPKMLKNKTRILVTHSMSYLPQVDVIV	840
Db	781	AVYSNADIYLFDDPLSAVDHVGKHI	FENVIGPKMLKNKTRILVTHSMSYLPQVDVIV	840
Qy	841	MSGGKISEMGSYQELLARQGAFAEFURTYAST	TEQEQDAEENGVTGVS	900
Db	841	MSGGKISEMGSYQELLARQGAFAEFURTYAST	TEQEQDAEENGVTGVS	900
Qy	901	LVTDSAGKQOROLSSSSVSGDISRHHNSTABL	QKAEAKKEETWKLMEADKAQGTGOVKL	960
Db	901	LVTDSAGKQOROLSSSSVSGDISRHHNSTABL	QKAEAKKEETWKLMEADKAQGTGOVKL	960
Qy	961	SVYWDYMKAIGLFISFLSIFLFCMCHVSA	LASNWLSLWTDDEI	1020
Db	961	SVYWDYMKAIGLFISFLSIFLFCMCHVSA	LASNWLSLWTDDEI	1020
Qy	1021	ALGISOGIAVFGYSMAVSIGGILASRCLH	VDLLHSILRSPMSFFERTPSGNLVNRFSKEL	1080
Db	1021	ALGISOGIAVFGYSMAVSIGGILASRCLH	VDLLHSILRSPMSFFERTPSGNLVNRFSKEL	1080
Qy	1081	DTVDSMIPVIKMFMSGSLFNVICACIVILLAT	PIAAIIIPPLGLIYFFVQRFVYASSRQL	1140
Db	1081	DTVDSMIPVIKMFMSGSLFNVICACIVILLAT	PIAAIIIPPLGLIYFFVQRFVYASSRQL	1140
Qy	1141	KRLSVSRSPVYSHFNETLLGVSVIRAFBEQ	ERFIHQSDLKVDENOKAYYPSIVANRWLA	1200
Db	1141	KRLSVSRSPVYSHFNETLLGVSVIRAFBEQ	ERFIHQSDLKVDENOKAYYPSIVANRWLA	1200
Qy	1201	VRLECVGNCIVLFAALFAVISRHSLSAGL	VGLSVSYSLQVTTYLNMLVRMSSEMETNIVA	1260
Db	1201	VRLECVGNCIVLFAALFAVISRHSLSAGL	VGLSVSYSLQVTTYLNMLVRMSSEMETNIVA	1260
Qy	1261	VERLKEYSETEKAPWQIOETAPPSSWPQGR	VEFRNYCLRYEEDLDFVLRHNVITNGG	1320
Db	1261	VERLKEYSETEKAPWQIOETAPPSSWPQGR	VEFRNYCLRYEEDLDFVLRHNVITNGG	1320
Qy	1321	EKVGIVGRTGAGKSSITLGLFRINESAEGE	IIIDGINIAKIGLHDLRFKTIIPQDPVLF	1380
Db	1321	EKVGIVGRTGAGKSSITLGLFRINESAEGE	IIIDGINIAKIGLHDLRFKTIIPQDPVLF	1380
Qy	1381	SGSLRNLNLPFGQSYDEEVWTSLELAHLKD	PFVSALPKLDHECAEGENLSVGORQLVCL	1440
Db	1381	SGSLRNLNLPFGQSYDEEVWTSLELAHLKD	PFVSALPKLDHECAEGENLSVGORQLVCL	1440
Qy	1441	ARALLRKTILVLDEATAVDLETDDLIOSTIR	TQFEDCTVLTIAHRLNTIMDYTRVIVL	1500
Db	1441	ARALLRKTILVLDEATAVDLETDDLIOSTIR	TQFEDCTVLTIAHRLNTIMDYTRVIVL	1500
Qy	1501	DKGEIOEGYCAPSLLQOQGLFYSMAXDAGLV	1531	
Db	1501	DKGEIOEGYCAPSLLQOQGLFYSMAXDAGLV	1531	
RESULT 3				
US-08-460-907B-4				
; Sequence 4, Application US/08460907B				
; Patent No. 5891724				
; GENERAL INFORMATION:				
; APPLICANT: Deeley, Roger G.				
; APPLICANT: Cole, Susan P.C.				
; TITLE OF INVENTION: METHODS FOR CONFERRING MULTIDRUG				
; TITLE OF INVENTION: RESISTANCE ON A CELL				
; NUMBER OF SEQUENCES: 9				
; CORRESPONDENCE ADDRESS:				
; ADDRESSEE: PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS				
; STREET: Queen's University at Kingston				
; CITY: Kingston				
; STATE: Ontario				
; COUNTRY: CANADA				
; ZIP: K7L 3N6				

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,907B
FILING DATE: 05-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923
FILING DATE: 27-OCT-1992
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Steeg, Carol Miernicki
REGISTRATION NUMBER: 39,539
REFERENCE/DOCKET NUMBER: Q1551
TELECOMMUNICATION INFORMATION:
TELEPHONE: (613) 545-2342
TELEFAX: (613) 545-6853
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1531 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-460-907B-4

```

```

Query Match      78.5%; Score 7860; DB 2; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MALRGFCADGSDPLMDNVTWNTSNPDFTKCFQNTLVWVPCFVLWACFPFFFLYLSRH 60
DB      1 MALRGFCADGSDPLMDNVTWNTSNPDFTKCFQNTLVWVPCFVLWACFPFFFLYLSRH 60

QY      61 DRGYTQMTPLNKTALGFLWIVCWADLFYSFWERSRGIFLAPVFLVSPITLLGITTLLA 120
DB      61 DRGYTQMTPLNKTALGFLWIVCWADLFYSFWERSRGIFLAPVFLVSPITLLGITTLLA 120

QY      121 TFLIQLERRKGVSQSGIMLTFWLVALCALAILRSKIMTALKEDAQVDLFRDITPYVYFS 180
DB      121 TFLIQLERRKGVSQSGIMLTFWLVALCALAILRSKIMTALKEDAQVDLFRDITPYVYFS 180

QY      181 LILLIQLVLSGDSRPLSETHDNPCESSASFLSRITFWITGLIVRGVROPLEGSD 240
DB      181 LILLIQLVLSGDSRPLSETHDNPCESSASFLSRITFWITGLIVRGVROPLEGSD 240

QY      241 LWSLNKEDTSEQVVPVLVKNKKCECAKTRKQPKVYVSSKDPAPKSSKYDANEEVEAL 300
DB      241 LWSLNKEDTSEQVVPVLVKNKKCECAKTRKQPKVYVSSKDPAPKSSKYDANEEVEAL 300

QY      301 IVKSPQKEWNSLFKVLKTTGPFYFLMSFFFKAIHDLNMFSGPQILKLLIKFVNDTKAPD 360
DB      301 IVKSPQKEWNSLFKVLKTTGPFYFLMSFFFKAIHDLNMFSGPQILKLLIKFVNDTKAPD 360

QY      361 WQGYFTVLLFVTAQLTVLHQVPHICFVSGMRKTAIVGAVYRKALVIINTSARKSSTV 420
DB      361 WQGYFTVLLFVTAQLTVLHQVPHICFVSGMRKTAIVGAVYRKALVIINTSARKSSTV 420

QY      421 GEIVNLSMVDQRFMDLATYINMIWSAPLOVILAYLLWNLGPSVLGAVAVMVLMPVN 480

```

```

DB      421 GEIVNLSMVDQRFMDLATYINMIWSAPLOVILAYLLWNLGPSVLGAVAVMVLMPVN 480
QY      481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKLYAWELAFKDKVLAIROBELKVLK 540
DB      481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKLYAWELAFKDKVLAIROBELKVLK 540
QY      541 KSAYLSAVGTFTWCTPFLVALCTPAVYVTTDENNILDQAOTAFVLSALFNILRFPFINILP 600
DB      541 KSAYLSAVGTFTWCTPFLVALCTPAVYVTTDENNILDQAOTAFVLSALFNILRFPFINILP 600
QY      601 MVISSIVQASVSLKELRIFLSHEELEPDSIERRPVKDGGGTSITVRNATTTWARSDEPT 660
DB      601 MVISSIVQASVSLKELRIFLSHEELEPDSIERRPVKDGGGTSITVRNATTTWARSDEPT 660
QY      661 LINGITFSIPEGALVAVVGVQCGKSSLLSALLAEMDKVEGHVAIKGSVAYVPPQAWIQND 720
DB      661 LINGITFSIPEGALVAVVGVQCGKSSLLSALLAEMDKVEGHVAIKGSVAYVPPQAWIQND 720
QY      721 SLRENILFGCQLEBPYRSVIOACALLPDLBILPSGDRTEIGEGVNLSSGQKQKVSUAR 780
DB      721 SLRENILFGCQLEBPYRSVIOACALLPDLBILPSGDRTEIGEGVNLSSGQKQKVSUAR 780
QY      781 AVYSNADIYLFDDPLSAVDHVGKHIPENVITGPKGMLKNKTRILVTHSMSYLPQVDVILV 840
DB      781 AVYSNADIYLFDDPLSAVDHVGKHIPENVITGPKGMLKNKTRILVTHSMSYLPQVDVILV 840
QY      841 MSGGKISEMGSYQELLARDGAFEFRTYASTEQEDAEENGVTGVSQPGKEAKQEMNGM 900
DB      841 MSGGKISEMGSYQELLARDGAFEFRTYASTEQEDAEENGVTGVSQPGKEAKQEMNGM 900
QY      901 LVTDGAGKQORQLSSSSSYSGDISRHNSNTEAELOKAEKEETWKLMEADKAOTGVK 960
DB      901 LVTDGAGKQORQLSSSSSYSGDISRHNSNTEAELOKAEKEETWKLMEADKAOTGVK 960
QY      961 SVYWDYMKAIGLFISFLSIFLMCNHVSALASNYWLSLTDPTVNGTOEHTKVRLSVYG 1020
DB      961 SVYWDYMKAIGLFISFLSIFLMCNHVSALASNYWLSLTDPTVNGTOEHTKVRLSVYG 1020
QY      1021 ALGISQGIAPVGYSMVAVSIGGILASRCLHVDLHLSILRSPMSFPFRTSGNLVNRFSKEL 1080
DB      1021 ALGISQGIAPVGYSMVAVSIGGILASRCLHVDLHLSILRSPMSFPFRTSGNLVNRFSKEL 1080
QY      1081 DTVDSDMPEVTKMPMGSLFNIVGACIVILLATPAAIIIPPLGLIYFFVQRFYVASSRQL 1140
DB      1081 DTVDSDMPEVTKMPMGSLFNIVGACIVILLATPAAIIIPPLGLIYFFVQRFYVASSRQL 1140
QY      1141 KRLESVSRSRPVYSHFNETLLGVSVIRAFEBQERFIHQSDLKVDENQKAYPSIVANRWLA 1200
DB      1141 KRLESVSRSRPVYSHFNETLLGVSVIRAFEBQERFIHQSDLKVDENQKAYPSIVANRWLA 1200
QY      1201 VRLECVGNCIVLFAALFAVISRHSLSAGLVCLSVSYSLQVTTYLNWLVRSMSMETNIVA 1260
DB      1201 VRLECVGNCIVLFAALFAVISRHSLSAGLVCLSVSYSLQVTTYLNWLVRSMSMETNIVA 1260
QY      1261 VERLKEYSETEKEAPWQIQETAPPSSWPQVGRVEFRNYCLRYREDLDFVLRHINVTNGG 1320
DB      1261 VERLKEYSETEKEAPWQIQETAPPSSWPQVGRVEFRNYCLRYREDLDFVLRHINVTNGG 1320
QY      1321 EKVGIVGRTGAGKSLTGLFRINESAEGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
DB      1321 EKVGIVGRTGAGKSLTGLFRINESAEGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
QY      1381 SGSLRWNLDFPSQYSDEEVTSLLEHLKDFVSALPKLDHCEAEGENISVGQRLVCL 1440
DB      1381 SGSLRWNLDFPSQYSDEEVTSLLEHLKDFVSALPKLDHCEAEGENISVGQRLVCL 1440
QY      1441 ARALLRKTILVDEATAAVDLETDLIQSTIRKQPEDCTVLTIAHRLNTIMDTRIVL 1500
DB      1441 ARALLRKTILVDEATAAVDLETDLIQSTIRKQPEDCTVLTIAHRLNTIMDTRIVL 1500
QY      1501 DKGEIOBYGAPSDLLQORGLFYSNAXDAGLV 1531
DB      1501 DKGEIOBYGAPSDLLQORGLFYSNAXDAGLV 1531

```

Db 1501 DKBEIOEYCAPSDLLQOQGLFYSMKADAGLV 1531

RESULT 4  
US-08-463-179A-4  
; Sequence 4, Application US/08463179A  
; Patent No. 6001563  
; GENERAL INFORMATION:  
; APPLICANT: Cole, Susan P.C.  
; APPLICANT: Delevy, Roger G.  
; TITLE OF INVENTION: METHODS FOR IDENTIFYING CHEMOSENSITIZERS  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street, suite 510  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/463,179A  
; FILING DATE:

CLASSIFICATION: 536  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: 07/966,923  
; FILING DATE: 27-OCT-1992  
; APPLICATION NUMBER: 08/029,340  
; FILING DATE: 8-MAR-1993  
; APPLICATION NUMBER: 08/141,893  
; FILING DATE: 26-OCT-1993  
; APPLICATION NUMBER: 08/407,207  
; FILING DATE: 20-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: DeConti, Giulio A. Jr.  
; REGISTRATION NUMBER: 31,503  
; REFERENCE/DOCKET NUMBER: PQ1-002CP8  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 227-5941  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1531 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-463-179A-4

Query Match 78.5%; Score 7860; DB 3; Length 1531;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGCSDGSDPLMDNMVNTWNTSNPDKCFQNTLVWVPCFYLMACFPFYLYLSRH 60  
Db 1 MALRGCSDGSDPLMDNMVNTWNTSNPDKCFQNTLVWVPCFYLMACFPFYLYLSRH 60  
QY 61 DRGYIQMTPLNKTALGELLMIWCVADLFYFWSRSGIFLAPVLPVSPDLLGITLLA 120  
Db 61 DRGYIQMTPLNKTALGELLMIWCVADLFYFWSRSGIFLAPVLPVSPDLLGITLLA 120  
QY 121 TELIQLERKGVQSSGIMLTFWLVALVLCALILRSKIMTALKEDAQVDLFRDITFYVPS 180  
Db 121 TELIQLERKGVQSSGIMLTFWLVALVLCALILRSKIMTALKEDAQVDLFRDITFYVPS 180  
QY 181 LLLIQLVLCFSDRSPFLSETIHDNPNCPSSASFLSRITFWITGLIVRGYRQPLEGSD 240  
Db 181 LLLIQLVLCFSDRSPFLSETIHDNPNCPSSASFLSRITFWITGLIVRGYRQPLEGSD 240  
QY 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPVVYSSKDPAPQPKSSKVDANEVEAL 300

Db 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPVVYSSKDPAPQPKSSKVDANEVEAL 300  
QY 301 IVKSPQKEWNPDLFKVLYKTGPGYFLMSFFPKAIHDLMMFSGPQILKLIKFVNDTKAPD 360  
Db 301 IVKSPQKEWNPDLFKVLYKTGPGYFLMSFFPKAIHDLMMFSGPQILKLIKFVNDTKAPD 360  
QY 361 WQGYFTYVLLFVTAQLQTLVLHQYFPHICFVSGMRIKTAVIGAVYRKALVITNSARKSSTV 420  
Db 361 WQGYFTYVLLFVTAQLQTLVLHQYFPHICFVSGMRIKTAVIGAVYRKALVITNSARKSSTV 420  
QY 421 GEIVNLSVDAQR FMDLATYINMISAPLOVLALYLWNLGSPVSLAGVAVMLVMPVN 480  
Db 421 GEIVNLSVDAQR FMDLATYINMISAPLOVLALYLWNLGSPVSLAGVAVMLVMPVN 480  
QY 481 AVMAKTKTYQVAHMKSKDNRIKLNWEILNGIKVLKYLAWELAFKDKVLAIRQEBELKVLK 540  
Db 481 AVMAKTKTYQVAHMKSKDNRIKLNWEILNGIKVLKYLAWELAFKDKVLAIRQEBELKVLK 540  
QY 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTTIDENNILDAQTAFVSLALFNILRFPNLILP 600  
Db 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTTIDENNILDAQTAFVSLALFNILRFPNLILP 600  
QY 601 MVISSIVQASVSLKELRIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWASDDPT 660  
Db 601 MVISSIVQASVSLKELRIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWASDDPT 660  
QY 661 LMGITTFIPEGALVAVGVQCGKSSLSALLAEMDKVEGHVAIKGSVAYVPPQQAWIQND 720  
Db 661 LMGITTFIPEGALVAVGVQCGKSSLSALLAEMDKVEGHVAIKGSVAYVPPQQAWIQND 720  
QY 721 SLRENILFGCOLLEPYRSVIAQACALLPDLEILPSGDRTEIGEKGVLNLSGGQKQVSLAR 780  
Db 721 SLRENILFGCOLLEPYRSVIAQACALLPDLEILPSGDRTEIGEKGVLNLSGGQKQVSLAR 780  
QY 781 AVYSNADIYLFDDPLSADAVHGVKHIPENVIGPKMKLNKTRILVTHSMYSYLPQVDVIV 840  
Db 781 AVYSNADIYLFDDPLSADAVHGVKHIPENVIGPKMKLNKTRILVTHSMYSYLPQVDVIV 840  
QY 841 MSGGKISEMGVYQELLARDGAPAEFLRTYASTEQEQDAEENGVTGSGPGKAKOMENGM 900  
Db 841 MSGGKISEMGVYQELLARDGAPAEFLRTYASTEQEQDAEENGVTGSGPGKAKOMENGM 900  
QY 901 LVTSAGKQLQRLSSSSSYSGDISRHHNSTAELQKAEAKKEETWKLMEADKAQGTQVKL 960  
Db 901 LVTSAGKQLQRLSSSSSYSGDISRHHNSTAELQKAEAKKEETWKLMEADKAQGTQVKL 960  
QY 961 SVYDYMKAIGLIFISFLIFPMCNHVSALASNYWLSLWTDPIVNGTQEHTKVRLSVYG 1020  
Db 961 SVYDYMKAIGLIFISFLIFPMCNHVSALASNYWLSLWTDPIVNGTQEHTKVRLSVYG 1020  
QY 1021 ALGISQGIATVFGYSMAVSIIGGILASRCLHVDLLHLSILRSPMSFFERTPSGNLVNRFSKEL 1080  
Db 1021 ALGISQGIATVFGYSMAVSIIGGILASRCLHVDLLHLSILRSPMSFFERTPSGNLVNRFSKEL 1080  
QY 1081 DTVDSMIEVIMKFMGSLFNIVIGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQL 1140  
Db 1081 DTVDSMIEVIMKFMGSLFNIVIGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQL 1140  
QY 1141 KRLESVSRSPVYSHNETLLGVSVIRAFEEQERFIHQSDLKVDENQKAYPSIVANRWLA 1200  
Db 1141 KRLESVSRSPVYSHNETLLGVSVIRAFEEQERFIHQSDLKVDENQKAYPSIVANRWLA 1200  
QY 1201 VRLECVGNICVILFAALFAVISRHSISAGLVGSYSIQVTTYLNLVWMSMETNIVA 1260  
Db 1201 VRLECVGNICVILFAALFAVISRHSISAGLVGSYSIQVTTYLNLVWMSMETNIVA 1260  
QY 1261 VERLKEYSETEKEAPWQIQETAPSSWPQVGRVEPRNYCLRYREDLDFVLRHINTYINGG 1320  
Db 1261 VERLKEYSETEKEAPWQIQETAPSSWPQVGRVEPRNYCLRYREDLDFVLRHINTYINGG 1320  
QY 1321 EKVGVIGRTGAGKSSLTGLFRINSAEGEIIDGINIAKGLHDLRPFKITTIPQDVLIF 1380

Db 1321 EKVGIVGRTGAGKSSLTGLPRINESAGEIIDIINIGINIAKIGLHDLRFKTIIPQDPVLV 1380  
 QY 1381 SGLRNLDLPFSQYSDDEEVTWTSLELAHLKDFVSALPDKLDHCECAGGENLSVGQRLVCL 1440  
 Db 1381 SGLRNLDLPFSQYSDDEEVTWTSLELAHLKDFVSALPDKLDHCECAGGENLSVGQRLVCL 1440  
 QY 1441 ARALLRKTKILVLDATAVDDTDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 Db 1441 ARALLRKTKILVLDATAVDDTDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 QY 1501 DKGEIQEYGAPSDLDLQOGLFYSMADAGLV 1531  
 Db 1501 DKGEIQEYGAPSDLDLQOGLFYSMADAGLV 1531

## RESULT 5

US-08-461-384B-4  
 ; Sequence 4, Application US/08461384B  
 ; Patent No. 6025473  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cole, Susan P.C.  
 ; APPLICANT: Deeley, Roger G.  
 ; TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS  
 ; NUMBER OF SEQUENCES: 10  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS  
 ; STREET: Queen's University at Kingston  
 ; CITY: Kingston  
 ; STATE: Ontario  
 ; COUNTRY: CANADA  
 ; ZIP: K7L 3N6  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: ASCII text  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/461,384B  
 ; FILING DATE: 05-JUN-95  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 07/966,923  
 ; FILING DATE: 27-OCT-1992  
 ; APPLICATION NUMBER: 08/029,340  
 ; FILING DATE: 8-MAR-1993  
 ; APPLICATION NUMBER: 08/141,893  
 ; FILING DATE: 26-OCT-1993  
 ; APPLICATION NUMBER: 08/407,207  
 ; FILING DATE: 20-MAR-1995  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Steeg, Carol Miernicki  
 ; REGISTRATION NUMBER: 39,539  
 ; REFERENCE/DOCKET NUMBER: Q1547  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (613) 545-2342  
 ; TELEFAX: (613) 545-6853  
 ; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1531 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 US-08-461-384B-4

Query Match 78.5%; Score 7860; DB 3; Length 1531;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MALRGFCADGSDPLWDMNVTWNTSNPDFTKCFQNTVLVWVPCFYLWACFPFFFLYLSRH 60  
 Db 1 MALRGFCADGSDPLWDMNVTWNTSNPDFTKCFQNTVLVWVPCFYLWACFPFFFLYLSRH 60  
 QY 61 DRGIQMTPLNKTALGFLFWVCWADLFYSFWERSRGIFLAPVFLVSPILLGITLLA 120

Db 61 DRGIQMTPLNKTALGFLFWVCWADLFYSFWERSRGIFLAPVFLVSPILLGITLLA 120  
 QY 121 TFLIQLERRKGVQSSGIMLTFWLVALVCALAILRSKIMTALKEDAQVDFRDITPYVYVFS 180  
 Db 121 TFLIQLERRKGVQSSGIMLTFWLVALVCALAILRSKIMTALKEDAQVDFRDITPYVYVFS 180  
 QY 181 LLLLTQLVLSCSFSDSPFLSETIHDNPNCPSSASPLSRITFWWITGLVIRGVQPLEGSD 240  
 Db 181 LLLLTQLVLSCSFSDSPFLSETIHDNPNCPSSASPLSRITFWWITGLVIRGVQPLEGSD 240  
 QY 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKVYSSKDPAPQKESKVDANEVEAL 300  
 Db 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKVYSSKDPAPQKESKVDANEVEAL 300  
 QY 301 IVKSPQKEWNPSPFLKVLKTYKTFPGPYFLMSFFPKAHDLMFSGPQILKLIIFVNDTKAPD 360  
 Db 301 IVKSPQKEWNPSPFLKVLKTYKTFPGPYFLMSFFPKAHDLMFSGPQILKLIIFVNDTKAPD 360  
 QY 361 WQGYFYTVLLFVTACLOTLVLHOYFHCFCVSGMRKTAIVIGAVYRKALVITNSARKSSTV 420  
 Db 361 WQGYFYTVLLFVTACLOTLVLHOYFHCFCVSGMRKTAIVIGAVYRKALVITNSARKSSTV 420  
 QY 421 GEIVNLMSVDAQRFMDLATYINMIWSAPLOQVILALYLMLNLGSPSVLAGVAVMLVPEVN 480  
 Db 421 GEIVNLMSVDAQRFMDLATYINMIWSAPLOQVILALYLMLNLGSPSVLAGVAVMLVPEVN 480  
 QY 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAWELAFKDKVLAIROBELKVLK 540  
 Db 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAWELAFKDKVLAIROBELKVLK 540  
 QY 541 KSAVLSAVGTFTWCTPFLVALCTFAVVTIDENNILDAQTAFVLSLAFNLIRPLNLTP 600  
 Db 541 KSAVLSAVGTFTWCTPFLVALCTFAVVTIDENNILDAQTAFVLSLAFNLIRPLNLTP 600  
 QY 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGSGTNSITVRNATFTWARSDDPT 660  
 Db 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGSGTNSITVRNATFTWARSDDPT 660  
 QY 661 LAGTTFSPREGALVAVVGVGCGKSSLLSALLAEMDKVEGHVAIKGSVAYVPOQAWTQND 720  
 Db 661 LAGTTFSPREGALVAVVGVGCGKSSLLSALLAEMDKVEGHVAIKGSVAYVPOQAWTQND 720  
 QY 721 SLRENILFGCOLEBPYYSRVIQACALLPDLLEILPSGDRTEIGEGVNLSSGQKQVSLAR 780  
 Db 721 SLRENILFGCOLEBPYYSRVIQACALLPDLLEILPSGDRTEIGEGVNLSSGQKQVSLAR 780  
 QY 781 AVYSNADIYLPDDPLSAVDAHVGKHI FENVIGPKMLKNKTRILVTHSMSYLPQVDVIV 840  
 Db 781 AVYSNADIYLPDDPLSAVDAHVGKHI FENVIGPKMLKNKTRILVTHSMSYLPQVDVIV 840  
 QY 841 MSGGKISEMGSYQELLARDGAFAEFLRTYASTEQEOQDAENGVTGVSQPGKEAKOMENG 900  
 Db 841 MSGGKISEMGSYQELLARDGAFAEFLRTYASTEQEOQDAENGVTGVSQPGKEAKOMENG 900  
 QY 901 LVTDGAKQLQORQLSSSSSYSGDISRHHNSTAELOKAEKKEETWKLMEADKAQTQGVKL 960  
 Db 901 LVTDGAKQLQORQLSSSSSYSGDISRHHNSTAELOKAEKKEETWKLMEADKAQTQGVKL 960  
 QY 961 SVYWDYMKAIGLFISFLSIFELFMCNHNVSALASNWLSTLTDPIVNGTQEHKTVRLSVYG 1020  
 Db 961 SVYWDYMKAIGLFISFLSIFELFMCNHNVSALASNWLSTLTDPIVNGTQEHKTVRLSVYG 1020  
 QY 1021 ALGISQGIAGVGYSMVSYIGIILASRCLHVDLHLSILRS PMSPPERTPSGNLVNRFSEKEL 1080  
 Db 1021 ALGISQGIAGVGYSMVSYIGIILASRCLHVDLHLSILRS PMSPPERTPSGNLVNRFSEKEL 1080  
 QY 1081 DTVDSDMPEVVKMFMGSLFNVI GACIVILLATPIAAIIIPPLGLIYFFVQFYVASSRQL 1140  
 Db 1081 DTVDSDMPEVVKMFMGSLFNVI GACIVILLATPIAAIIIPPLGLIYFFVQFYVASSRQL 1140  
 QY 1141 KRLESVSRSPVYSHFNETLLGVSVIRAFEBQERFIHOSDLKVDENQKAYVPSIVANRWLA 1200  
 Db 1141 KRLESVSRSPVYSHFNETLLGVSVIRAFEBQERFIHOSDLKVDENQKAYVPSIVANRWLA 1200

Qy	1201	VRLECVGNCIVLFAALFAVTSRHSLSAGLVGLSVSYSLQVTTVTLNMLVRMSSEMTNIVA	1260
Db	1201	VRLECVGNCIVLFAALFAVTSRHSLSAGLVGLSVSYSLQVTTVTLNMLVRMSSEMTNIVA	1260
Qy	1261	VERLKEYSETEKAPWQIQETAPSSPWPQGRVEFRNYCLRYREDLDFVLRHINVTINGG	1320
Db	1261	VERLKEYSETEKAPWQIQETAPSSPWPQGRVEFRNYCLRYREDLDFVLRHINVTINGG	1320
Qy	1321	EKVGIVGRTGAGKSSLTGLFRINESAEGSIIIDGINIAKIGHLDLRFKTIITIPQDPVLV	1380
Db	1321	EKVGIVGRTGAGKSSLTGLFRINESAEGSIIIDGINIAKIGHLDLRFKTIITIPQDPVLV	1380
Qy	1381	SGSLRWMLDPFSQYSDEEVTSLLEAHLKDFVSALPKLDHECAEGGENLUSVQORQLVCL	1440
Db	1381	SGSLRWMLDPFSQYSDEEVTSLLEAHLKDFVSALPKLDHECAEGGENLUSVQORQLVCL	1440
Qy	1441	ARALLRKTILVLDEATAAVDLETDLLIQSTIRTQFEDCTVLVLTIAHRLNTIMDYTRVIVL	1500
Db	1441	ARALLRKTILVLDEATAAVDLETDLLIQSTIRTQFEDCTVLVLTIAHRLNTIMDYTRVIVL	1500
Qy	1501	DKGEIQBYGAPSDLLQORGLFYSMWAKDAGLV	1531
Db	1501	DKGEIQBYGAPSDLLQORGLFYSMWAKDAGLV	1531

RESULT 6  
US-09-647-140B-19  
; Sequence 19, Application US/09647140B  
; Patent No. 6803184  
; GENERAL INFORMATION:  
; APPLICANT: Fox Chase Cancer Center  
; APPLICANT: Kruh, Gary D.  
; APPLICANT: Lee, Kun  
; APPLICANT: Belinsky, Martin G.  
; APPLICANT: Bain, Lisa J.  
; TITLE OF INVENTION: MRP-Related ABC Transporter Encoding  
; FILE OF INVENTION: Nucleic Acids and Methods of Use Thereof  
; TITLE REFERENCE: FCCC 98-02  
; CURRENT APPLICATION NUMBER: US/09/647,140B  
; CURRENT FILING DATE: 2001-05-21  
; PRIOR APPLICATION NUMBER: PCT/US99/06644  
; PRIOR FILING DATE: 1999-03-26  
; PRIOR APPLICATION NUMBER: 60/079,759  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/095,153  
; PRIOR FILING DATE: 1998-08-03  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 19  
; LENGTH: 1531  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-09-647-140B-19

	Query Match	78.5%;	Score 7860;	DB 4;	Length 1531;
	Best Local Similarity	100.0%;	Prod. No. 0;		
	Matches 1531;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MALRGFCSADGSDPLDWNVTWNTNSDPDKCFQNTVLVWVPCFYLWACPPFFYLYLSRH	60		
Db	1	MALRGFCSADGSDPLDWNVTWNTNSDPDKCFQNTVLVWVPCFYLWACPPFFYLYLSRH	60		
Qy	61	DRGYQMTPLNKTKTALGFLLIWVCWADLPFSFWERSRGIFLAPVLSVSPTLIGITITLLA	120		
Db	61	DRGYQMTPLNKTKTALGFLLIWVCWADLPFSFWERSRGIFLAPVLSVSPTLIGITITLLA	120		
Qy	121	TFLIQLERRKGVQSSGIMLTFWLVALVCALAILRSKIMTKALKEDAQVDLFRDITFFYYFS	180		
Db	121	TFLIQLERRKGVQSSGIMLTFWLVALVCALAILRSKIMTKALKEDAQVDLFRDITFFYYFS	180		
Qy	181	LLATQLVLSCFSDRSPIFSETIHPNPNCPSSASFLSRITFWWTITGLIVRGYROPLEGSD	240		

181	LLLIQLVLSCFSDRSLPSETHDNPNCPESSASFSLRITFWITGLIVRGYRQPLEGSD	240
241	LMSLNKEDTSEQVVPVLVRNWKKECAKTRKQPVKVYVYSSKPAQPKESSKVDANEVEAL	300
241	LMSLNKEDTSEQVVPVLVRNWKKECAKTRKQPVKVYVYSSKPAQPKESSKVDANEVEAL	300
301	IVKSPOKENWPSLFLKVLKXKTFGYPFLMSFFPKAIHDLMMPSGQIILKLIKFVNDTKAPD	360
301	IVKSPOKENWPSLFLKVLKXKTFGYPFLMSFFPKAIHDLMMPSGQIILKLIKFVNDTKAPD	360
361	WQGYFYTVLLFVTVACLOTLVLHQYFHCFSVGMRIKTAVIGAVYRKALVTITNSARKSSTV	420
361	WQGYFYTVLLFVTVACLOTLVLHQYFHCFSVGMRIKTAVIGAVYRKALVTITNSARKSSTV	420
421	GEIVNLMSDAQRFMDLATYINMWSAPLQVILALYLLMLNLGSPSLAGVAVMVLMPVFN	480
421	GEIVNLMSDAQRFMDLATYINMWSAPLQVILALYLLMLNLGSPSLAGVAVMVLMPVFN	480
481	AVNAMTKTYQVAHMKSKONRIKLMEILINGIKVLKYAWELAFKQKVLAIROEELKVLK	540
481	AVNAMTKTYQVAHMKSKONRIKLMEILINGIKVLKYAWELAFKQKVLAIROEELKVLK	540
541	KSAYLSAVGFTWVCTPELVALCTEAVYVTIDENNILDAQTAFVSLALFNILFPNLNLP	600
541	KSAYLSAVGFTWVCTPELVALCTEAVYVTIDENNILDAQTAFVSLALFNILFPNLNLP	600
601	MWISSVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGGTNSITVRNATFTWARSDDPT	660
601	MWISSVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGGTNSITVRNATFTWARSDDPT	660
661	LNGITPSIEGALVAVVGQVGCCKSSILLSALLAEMDKVEGHVAIKGSVAVTPOQAWIQND	720
661	LNGITPSIEGALVAVVGQVGCCKSSILLSALLAEMDKVEGHVAIKGSVAVTPOQAWIQND	720
721	SURENITLFGCOLLEEPYRSVIOACALLPDLLEILPDSGRTEIGBKGNLSGGQKQKVSAR	780
721	SURENITLFGCOLLEEPYRSVIOACALLPDLLEILPDSGRTEIGBKGNLSGGQKQKVSAR	780
781	AVYSNADIYLFDDPLSAVDHVGKHFENYVIGPKGMLKNKTRILVTHSMSYLPQVDVIIV	840
781	AVYSNADIYLFDDPLSAVDHVGKHFENYVIGPKGMLKNKTRILVTHSMSYLPQVDVIIV	840
841	MSGGKISEMGSYQELLARDGCAFAEFLRTVASTEQEQDAEENGVTGSGPQKEAKQWENG	900
841	MSGGKISEMGSYQELLARDGCAFAEFLRTVASTEQEQDAEENGVTGSGPQKEAKQWENG	900
901	LVTDSAGKOLORLSSSSVSGDISRHNSSTAELOKAEAKKEETWKLMEADKAQTGVOKL	960
901	LVTDSAGKOLORLSSSSVSGDISRHNSSTAELOKAEAKKEETWKLMEADKAQTGVOKL	960
961	SVTWDTYKAIGLFISFLSIFLFCMCHVSALASNYWLSLWTDDEIVNGTQEHKTVRLSVYG	1020
961	SVTWDTYKAIGLFISFLSIFLFCMCHVSALASNYWLSLWTDDEIVNGTQEHKTVRLSVYG	1020
1021	ALGISOGIAVFGYSMAVSGGIIASRCLHVDLLHSLRSPMSFPERTPSGNLVNRFSEKEL	1080
1021	ALGISOGIAVFGYSMAVSGGIIASRCLHVDLLHSLRSPMSFPERTPSGNLVNRFSEKEL	1080
1081	DTVDSMIPEVIMKMGSLFNVTGACIVILLATPIAAIITPPLGLIYFFVQREYVASSRQL	1140
1081	DTVDSMIPEVIMKMGSLFNVTGACIVILLATPIAAIITPPLGLIYFFVQREYVASSRQL	1140
1141	KRLSVSRSPVYSHFNETLLGVSVIRAFEQEERFIHQSDLKVDENQKAYYPSIVANRWLA	1200
1141	KRLSVSRSPVYSHFNETLLGVSVIRAFEQEERFIHQSDLKVDENQKAYYPSIVANRWLA	1200
1201	VRLECVGNCIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTVNLMLVRMSSSEMETHIVA	1260
1201	VRLECVGNCIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTVNLMLVRMSSSEMETHIVA	1260
1261	VERLKEYSETEKAPHQIQETAPSSWPQVGRVEFRNYCLRYREDLDFVLRHINVTINGG	1320
1261	VERLKEYSETEKAPHQIQETAPSSWPQVGRVEFRNYCLRYREDLDFVLRHINVTINGG	1320

QY 1321 EKVGIVGRTGAGKSSLTGLFRINESAGEBIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
 DB 1321 EKVGIVGRTGAGKSSLTGLFRINESAGEBIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
 QY 1381 SGLRNLNLPFFSQYSDDEEVMVTSLELAHLKDFVSALPDKLHCEAEGENLSVGQRLVCL 1440  
 DB 1381 SGLRNLNLPFFSQYSDDEEVMVTSLELAHLKDFVSALPDKLHCEAEGENLSVGQRLVCL 1440  
 QY 1441 ARALRKTILVLDEATAVDLETDLIQSTIRTPQEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 DB 1441 ARALRKTILVLDEATAVDLETDLIQSTIRTPQEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 QY 1501 DKGEIOEYGAPSDLLQORGLFYSMADAGLV 1531  
 DB 1501 DKGEIOEYGAPSDLLQORGLFYSMADAGLV 1531

RESULT 7  
 US-08-141-893-2  
 ; Sequence 2, Application US/08141893  
 ; Patent No. 5489519  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cole, Susan P.C.  
 ; TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEIN  
 ; NUMBER OF SEQUENCES: 2  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: LAHIVE & COCKFIELD  
 ; STREET: 60 STATE STREET, SUITE 510  
 ; CITY: BOSTON  
 ; STATE: MASSACHUSETTS  
 ; COUNTRY: USA  
 ; ZIP: 02109  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/141,893  
 ; FILING DATE:  
 ; CLASSIFICATION: 424  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 07/966,923; 08/029,340  
 ; FILING DATE: 27-OCT-1992; 8-MAR-1993  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: DeConti, Giulio A. Jr.  
 ; REGISTRATION NUMBER: 31,503  
 ; REFERENCE/DOCKET NUMBER: PQI-002  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (617) 227-7400  
 ; TELEFAX: (617) 227-5149  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1531 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 US-08-141-893-2

Query Match 78.4%; Score 7849; DB 1; Length 1531;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 MALRGFCSDGSDPLNDWNNVTNTSNPDFTKCFQNTVLVWVPCFYLWACFPFFLYLSRH 60  
 DB 1 MALRGFCSDGSDPLNDWNNVTNTSNPDFTKCFQNTVLVWVPCFYLWACFPFFLYLSRH 60  
 QY 61 DRGIQMTPLNKTALGFLWIVCWADLFYSFWRSGRGIPLAVFLVSPFLIGITLLA 120  
 DB 61 DRGIQMTPLNKTALGFLWIVCWADLFYSFWRSGRGIPLAVFLVSPFLIGITLLA 120

QY 121 TFLIQLERRKGVSSGIMLTFWLVALVCAIALILRSKIMTALKEDAQVDLFRDITFYVYFS 180  
 DB 121 TFLIQLERRKGVSSGIMLTFWLVALVCAIALILRSKIMTALKEDAQVDLFRDITFYVYFS 180  
 QY 181 LLLIQLVLSGSDRPLSFSETHDNPCESSASFLSRITFWITGLIVRGVQRPLEGSD 240  
 DB 181 LLLIQLVLSGSDRPLSFSETHDNPCESSASFLSRITFWITGLIVRGVQRPLEGSD 240  
 QY 241 LWSLNKEDTSQGVVPLVKNWKECAKTRKQPVKVYSSKDPAPKESKVDANEVEAL 300  
 DB 241 LWSLNKEDTSQGVVPLVKNWKECAKTRKQPVKVYSSKDPAPKESKVDANEVEAL 300  
 QY 301 IVKSPQKEWNSLKFVLYKTFPGPYFLMSFFPKAHLDMFSGPQILKLLIKFVNDTKAPD 360  
 DB 301 IVKSPQKEWNSLKFVLYKTFPGPYFLMSFFPKAHLDMFSGPQILKLLIKFVNDTKAPD 360  
 QY 361 WQGYFYTVLLFVTACLOTVLHQLVHCFVSGMRIKTAIVIGAVYRKALVITNSARKSSTV 420  
 DB 361 WQGYFYTVLLFVTACLOTVLHQLVHCFVSGMRIKTAIVIGAVYRKALVITNSARKSSTV 420  
 QY 421 GEIVNLMSVDAQRFMDLATYINMIWSAPLOVILALYLLWNLGSPSVLAGVAVMLWVPVN 480  
 DB 421 GEIVNLMSVDAQRFMDLATYINMIWSAPLOVILALYLLWNLGSPSVLAGVAVMLWVPVN 480  
 QY 481 AVMAKTKTYOVAAHMKSKDNRIKLMNEILNGIKVLYKLYAWELAFKDKVLAIRQELKVLK 540  
 DB 481 AVMAKTKTYOVAAHMKSKDNRIKLMNEILNGIKVLYKLYAWELAFKDKVLAIRQELKVLK 540  
 QY 541 KSAYLSAVGTFTWCTPFLVALCTFAVVVTTIDENNILDAQTAFAVLSLAFNLIRPLNLIP 600  
 DB 541 KSAYLSAVGTFTWCTPFLVALCTFAVVVTTIDENNILDAQTAFAVLSLAFNLIRPLNLIP 600  
 QY 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIRRRPVKGGGTNSITVRNATFTWARSDDPT 660  
 DB 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIRRRPVKGGGTNSITVRNATFTWARSDDPT 660  
 QY 661 LNGTHSIPEGALVAVVGVGCGKSLLSALLAEMDKVEGHVALKGSVAVVPOQAWTQND 720  
 DB 661 LNGTHSIPEGALVAVVGVGCGKSLLSALLAEMDKVEGHVALKGSVAVVPOQAWTQND 720  
 QY 721 SLRENILFGQLEBPYYRSVIQACALLPDLLEILPSGDRTEIGEKNVLSGQKQKRVSLAR 780  
 DB 721 SLRENILFGQLEBPYYRSVIQACALLPDLLEILPSGDRTEIGEKNVLSGQKQKRVSLAR 780  
 QY 781 AVYSNADIYLFDDPLSAVDHVGKHI FENVIGPKGMLKNKTRILVITHSMSVLPQVDVILV 840  
 DB 781 AVYSNADIYLFDDPLSAVDHVGKHI FENVIGPKGMLKNKTRILVITHSMSVLPQVDVILV 840  
 QY 841 MSGGKISEMGSYQELLARDGAFABFLRTYASTEQDAEENGVTGVSQPGKEAKQEMNGM 900  
 DB 841 MSGGKISEMGSYQELLARDGAFABFLRTYASTEQDAEENGVTGVSQPGKEAKQEMNGM 900  
 QY 901 LVTDGAKQLQRLSSSSSYSGDI SRHNSHSTAELQKAEKKEETWKLMEADKAQTQGVKL 960  
 DB 901 LVTDGAKQLQRLSSSSSYSGDI SRHNSHSTAELQKAEKKEETWKLMEADKAQTQGVKL 960  
 QY 961 SVYDYMKAIGLFI SFLSI FLFNCNHVSALASNYLSLWTDTPVNGTQEHKTVLSVYG 1020  
 DB 961 SVYDYMKAIGLFI SFLSI FLFNCNHVSALASNYLSLWTDTPVNGTQEHKTVLSVYG 1020  
 QY 1021 ALGISQGIAGVFGYSMAVSIIGGILASRCLHVDLHLSILRSPMSFFERTPSGNLVNRFSEKEL 1080  
 DB 1021 ALGISQGIAGVFGYSMAVSIIGGILASRCLHVDLHLSILRSPMSFFERTPSGNLVNRFSEKEL 1080  
 QY 1081 DTVDSMIPEVIKMFMSGLFNVI GACIVILLATPIAAIIIPPLGLIYFFVQRFVYASSRQL 1140  
 DB 1081 DTVDSMIPEVIKMFMSGLFNVI GACIVILLATPIAAIIIPPLGLIYFFVQRFVYASSRQL 1140  
 QY 1141 KRLESVRSRSPVYSHENETLLGVSVIRAFEBQERFIHQSDLKVDENOKAVYPSIVANRWLA 1200  
 DB 1141 KRLESVRSRSPVYSHENETLLGVSVIRAFEBQERFIHQSDLKVDENOKAVYPSIVANRWLA 1200  
 QY 1201 VRLECVGNCIVLFAALFAVISHSLSAGLVLSVSYSLQVTTYLNMLVNRMSSEMETNIVA 1260

1201 VRLECGNCVILFAALFAVISHSLAGLVLSVSYSLQVTTYNLWLRMSSEMETNIVA 1260  
1261 VERLKEYSETEKAPQIOETAPSSWPQVGRVFRNYCLRVREDLDVFLRHINVTINGG 1320  
1261 VERLKEYSETEKAPQIOETREPPSSWPQVGRVFRNYCLRVREDLDVFLRHINVTINGG 1320  
1321 EKVGIVGRTGAGKSSITLGLFRINESAEGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
1321 EKVGIVGRTGAGKSSITLGLFRINESAEGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
1381 SGLRNWLPFPQSDEEYVWTSLELAHLKDFVSALPDKLDHECABEGENLSVGQRLVCL 1440  
1381 SGLRNWLPFPQSDEEYVWTSLELAHLKDFVSALPDKLDHECABEGENLSVGQRLVCL 1440  
1441 ARALLRKTILVLDATAVDLETDLIQSTIRTFQEDCTVLTIAHRLNTIMDYTRVIVL 1500  
1441 ARALLRKTILVLDATAVDLETDLIQSTIRTFQEDCTVLTIAHRLNTIMDYTRVIVL 1500  
1501 DKGEIOEYGAPSDLLQQRGLFYSMAXDAGLV 1531  
1501 DKGEIOEYGAPSDLLQQRGLFYSMAXDAGLV 1531

RESULT 8

US-08-463-092B-2  
Sequence 2, Application US/08463092B  
Patent No. 5766880  
GENERAL INFORMATION:  
APPLICANT: Cole, Susan P.C.  
APPLICANT: Dealey, Roger G.  
TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ENCODING  
TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESS: PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS  
STREET: Queen's University at Kingston  
CITY: Kingston  
STATE: Ontario  
COUNTRY: CANADA  
ZIP: K7L 3N6  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/463,092B  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/966,923  
FILING DATE: 27-OCT-1992  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/029,340  
FILING DATE: 8-MAR-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/141,893  
FILING DATE: 26-OCT-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/407,207  
FILING DATE: 20-MAR-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Steeg, Carol Miernicki  
REGISTRATION NUMBER: 39,539  
REFERENCE/DOCKET NUMBER: Q1546  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (613) 545-2342  
TELEFAX: (613) 545-6853

INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1531 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-463-092B-2  
Query Match 78.4%; Score 7849; DB 1; Length 1531;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 MALRGCSDGSDPLMDNNTNTNPNPTKCFQNTLVVWPCFYLVWPCFYLYLSRH 60  
DB 1 MALRGCSDGSDPLMDNNTNTNPNPTKCFQNTLVVWPCFYLVWPCFYLYLSRH 60  
QY 61 DRGYTQMTPLNKTALGFLMWVWADLFYFWSRSRGIFLAPVFLVSPILLGITLLA 120  
DB 61 DRGYTQMTPLNKTALGFLMWVWADLFYFWSRSRGIFLAPVFLVSPILLGITLLA 120  
QY 121 TFLIQLERRKGVQSSGIMLTFWLVVALCALALIRSKIMTALKEDAQVDLFRDITFVYFS 180  
DB 121 TFLIQLERRKGVQSSGIMLTFWLVVALCALALIRSKIMTALKEDAQVDLFRDITFVYFS 180  
QY 181 LLLIQLVLSCFSDRSPLETFSETIHDNPNCPSSASFLSRITFWMITGLIVRGYRQPLEGSD 240  
DB 181 LLLIQLVLSCFSDRSPLETFSETIHDNPNCPSSASFLSRITFWMITGLIVRGYRQPLEGSD 240  
QY 241 LWSLNKEDTSEQVVPVVKNNWKECAKTRKQPVKVYSSKDPAPKESKVDANEVEAL 300  
DB 241 LWSLNKEDTSEQVVPVVKNNWKECAKTRKQPVKVYSSKDPAPKESKVDANEVEAL 300  
QY 301 IVKSPQKEWNPFLFKVLYKTFPGYFLMSPFFPKAIHDLMMFSGPOLIKLIKFVNDTKAPD 360  
DB 301 IVKSPQKEWNPFLFKVLYKTFPGYFLMSPFFPKAIHDLMMFSGPOLIKLIKFVNDTKAPD 360  
QY 361 WQGYFYTVLLFVTACLTQLVLHQYFHCIFVSGMRKTAIVAGYRKAALVITNSARKSSTV 420  
DB 361 WQGYFYTVLLFVTACLTQLVLHQYFHCIFVSGMRKTAIVAGYRKAALVITNSARKSSTV 420  
QY 421 GEIVNLSVDAQRFDMLATYINMIWSAPLOVTLALYLLNLNLPVSVAGVAVMLVMPVN 480  
DB 421 GEIVNLSVDAQRFDMLATYINMIWSAPLOVTLALYLLNLNLPVSVAGVAVMLVMPVN 480  
QY 481 AVAMKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKLYAWELAFKDKVLAIRQEEELKVLK 540  
DB 481 AVAMKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKLYAWELAFKDKVLAIRQEEELKVLK 540  
QY 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTDENNILDAQTAFVSLALFNILRFPPLNLP 600  
DB 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTDENNILDAQTAFVSLALFNILRFPPLNLP 600  
QY 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERPPVKDGGTNSITVRNATFTWASDDPT 660  
DB 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERPPVKDGGTNSITVRNATFTWASDDPT 660  
QY 661 LMGITFSPGALVAVGVQCGKSLLSALLAEMDKVEGHVAIKGSVAVYVQQAQWIND 720  
DB 661 LMGITFSPGALVAVGVQCGKSLLSALLAEMDKVEGHVAIKGSVAVYVQQAQWIND 720  
QY 721 SLRENILFGCQLEEPYRSVIOACALLPDLPLPSGDRTEIGEKGVLNSGGQKQVSLAR 780  
DB 721 SLRENILFGCQLEEPYRSVIOACALLPDLPLPSGDRTEIGEKGVLNSGGQKQVSLAR 780  
QY 781 AVYSNADIYLFDDPLSAVDHVGKHI FENVIGPKMLKNKTRILVTHSHSYLPQVDVIV 840  
DB 781 AVYSNADIYLFDDPLSAVDHVGKHI FENVIGPKMLKNKTRILVTHSHSYLPQVDVIV 840  
QY 841 MSGGKISEMGSYQELLARDGAFABFLRTYASTPEQSDAEENGVTGSGPKAKQWENGM 900  
DB 841 MSGGKISEMGSYQELLARDGAFABFLRTYASTPEQSDAEENGVTGSGPKAKQWENGM 900  
QY 901 LVTDSAGKQQLSLSSSSSSYSGDI SRHHNSTAELQKAEAKKEETWKLMEADKAQTQGVKL 960

Db 901 LVTDGAGLQRLSSSSSYSGDISRHHNSTAEKAEAKKEETWKLMEADKAQTQGVKL 960  
QY 961 SVYDYMKAIGLFLSFLFNCNHNVSALSNYLSLWTDPIVNGTOEHTKVRLSVYG 1020  
Db 961 SVYDYMKAIGLFLSFLFNCNHNVSALSNYLSLWTDPIVNGTOEHTKVRLSVYG 1020  
QY 1021 ALGISQIAVFGYSMAVSGGILASRLHVDLHLSILRSPMSFFERTPSGNLVNRSKEL 1080  
Db 1021 ALGISQIAVFGYSMAVSGGILASRLHVDLHLSILRSPMSFFERTPSGNLVNRSKEL 1080  
QY 1081 DTVDSEMPVIMFMGSLFNIVGACIVILLATPIAAIIIPPLGLIYFFVQRFVASSRQL 1140  
Db 1081 DTVDSEMPVIMFMGSLFNIVGACIVILLATPIAAIIIPPLGLIYFFVQRFVASSRQL 1140  
QY 1141 KRLESVSRSPVSHNETLLGVSIVIRAFEEQERFIHQSDLKVDENOKAYPPIVANNWLA 1200  
Db 1141 KRLESVSRSPVSHNETLLGVSIVIRAFEEQERFIHQSDLKVDENOKAYPPIVANNWLA 1200  
QY 1201 VRLECGNCIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTVLNWLVRMSSEMETNIVA 1260  
Db 1201 VRLECGNCIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTVLNWLVRMSSEMETNIVA 1260  
QY 1261 VERLKEYSETEKAPWQIQETAPPSWPQGRVFEFRNYCLRYREDLDFVLRHINVITNGG 1320  
Db 1261 VERLKEYSETEKAPWQIQETAPPSWPQGRVFEFRNYCLRYREDLDFVLRHINVITNGG 1320  
QY 1321 EKVGIVGRTGAGKSSITLGLFRINSAEGEIIIDGINIAKIGHDLRFKTIIPQDPVLF 1380  
Db 1321 EKVGIVGRTGAGKSSITLGLFRINSAEGEIIIDGINIAKIGHDLRFKTIIPQDPVLF 1380  
QY 1381 SGLSRMLNLPFFQYSDDEEVMTSLELAHLKDFVSALPKDLHCEAGENLSYGQRLVCL 1440  
Db 1381 SGLSRMLNLPFFQYSDDEEVMTSLELAHLKDFVSALPKDLHCEAGENLSYGQRLVCL 1440  
QY 1441 ABALLARKTILVLDENATAVDLETDLIQSTTRTQFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
Db 1441 ABALLARKTILVLDENATAVDLETDLIQSTTRTQFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
QY 1501 DKGEIQEYCAPSDLLQORGLFYSMKADAGLV 1531  
Db 1501 DKGEIQEYCAPSDLLQORGLFYSMKADAGLV 1531

## RESULT 9

US-08-462-109A-2  
; Sequence 2, Application US/08462109A  
; Patent No. 582875  
; GENERAL INFORMATION:  
; APPLICANT: Cole, Susan P.C.  
; APPLICANT: Deeley, Roger G.  
; TITLE OF INVENTION: METHODS FOR IDENTIFYING  
; TITLE OF INVENTION: MULTIDRUG RESISTANT TUMOR CELLS  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street, suite 510  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/462.109A  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/966,923  
; FILING DATE: 27-OCT-1992

; APPLICATION NUMBER: 08/029,340  
; FILING DATE: 8-MAR-1993  
; APPLICATION NUMBER: 08/141,893  
; FILING DATE: 26-OCT-1993  
; APPLICATION NUMBER: 08/407,207  
; FILING DATE: 20-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: DeConti, Giulio A. Jr.  
; REGISTRATION NUMBER: 31,503  
; REFERENCE/DOCKET NUMBER: PQI-002CP4  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 227-5941  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1531 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-462-109A-2  
Query Match 78.4%; Score 7849; DB 2; Length 1531;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 MALRGFCSDGSDPLDWNVTNTSNPDFTKCFQNTVLVWVPCFYLWACFPFFLYLSRH 60  
Db 1 MALRGFCSDGSDPLDWNVTNTSNPDFTKCFQNTVLVWVPCFYLWACFPFFLYLSRH 60  
QY 61 DRGVIQMTPLNKTALGFLWIVCWADLFYSFWERSRGIFLAPVFLVSPITLLGITTLLA 120  
Db 61 DRGVIQMTPLNKTALGFLWIVCWADLFYSFWERSRGIFLAPVFLVSPITLLGITTLLA 120  
QY 121 TFLIQLERKGVQSSGIMLTFLWALVALCALAILRSKIMTALKEDAQVDLFRDITFYVYFS 180  
Db 121 TFLIQLERKGVQSSGIMLTFLWALVALCALAILRSKIMTALKEDAQVDLFRDITFYVYFS 180  
QY 181 LLLIQLVLSCFSDRSPFLSETIHDNPNCPSSASFLSRITFWMTGLIVRGVROPLEGSD 240  
Db 181 LLLIQLVLSCFSDRSPFLSETIHDNPNCPSSASFLSRITFWMTGLIVRGVROPLEGSD 240  
QY 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKVYSSKDPAPQKSSKVDANEVEAL 300  
Db 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKVYSSKDPAPQKSSKVDANEVEAL 300  
QY 301 IVKSPQKEWNPFLFKVLYKTGPGPYFLMSFPFKAIHDLMMFSGPOILKLLIKFVNDTKAPD 360  
Db 301 IVKSPQKEWNPFLFKVLYKTGPGPYFLMSFPFKAIHDLMMFSGPOILKLLIKFVNDTKAPD 360  
QY 361 WQGYFYTVLLFVTACLOTVLVHOVFHICFVSGMRIKTAVIGAVYRKALVITNSARKSSTV 420  
Db 361 WQGYFYTVLLFVTACLOTVLVHOVFHICFVSGMRIKTAVIGAVYRKALVITNSARKSSTV 420  
QY 421 GEIVNLMSVDAQREFMDLATYINMIWSAPLOVILALYLLNMLGPSVLGAVVWMLVMPVN 480  
Db 421 GEIVNLMSVDAQREFMDLATYINMIWSAPLOVILALYLLNMLGPSVLGAVVWMLVMPVN 480  
QY 481 AVMAKTKTYQVAHMKSKDNRIKLMBEILNGIKVLYAWELAFKDLVLAIRQBELKVLK 540  
Db 481 AVMAKTKTYQVAHMKSKDNRIKLMBEILNGIKVLYAWELAFKDLVLAIRQBELKVLK 540  
QY 541 KSAYLSAVGTFTWVCTPFLVALCTFAVVVITDENNILDQAOTAFVSLAFNLFRPLNLLP 600  
Db 541 KSAYLSAVGTFTWVCTPFLVALCTFAVVVITDENNILDQAOTAFVSLAFNLFRPLNLLP 600  
QY 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGSGTNSITVRNATFTWARSDDPT 660  
Db 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGSGTNSITVRNATFTWARSDDPT 660  
QY 661 LINGITFSIPEGALVAVVGVGCGKSSLLSALLAEMDKVEGHVAIKGSVAVVPOQAWTQND 720  
Db 661 LINGITFSIPEGALVAVVGVGCGKSSLLSALLAEMDKVEGHVAIKGSVAVVPOQAWTQND 720

```

QY 721 SURENIFGCGLEEBYRSVIOACALLPDLEILPSGDRTEIGEKGVLNLSGGQKQVSLAR 780
Db 721 SURENIFGCGLEEBYRSVIOACALLPDLEILPSGDRTEIGEKGVLNLSGGQKQVSLAR 780
QY 781 AVYSNADILYFDDPLSADVAHVGHIFENVIGPKGMLKKNKTRILVTHSMSYLPQVDVLIIV 840
Db 781 AVYSNADILYFDDPLSADVAHVGHIFENVIGPKGMLKKNKTRILVTHSMSYLPQVDVLIIV 840
QY 841 MSGGKISEMGYSQELLARDGAFELRTVYASTEQEQDAEENGVTGVSQPGKEAKOMNGM 900
Db 841 MSGGKISEMGYSQELLARDGAFELRTVYASTEQEQDAEENGVTGVSQPGKEAKOMNGM 900
QY 901 LVTDGAGQLOLQSSSSYSYSDISRHNSTAELOKAEAKKEETWKLMEADKAQGTQVVL 960
Db 901 LVTDGAGQLOLQSSSSYSYSDISRHNSTAELOKAEAKKEETWKLMEADKAQGTQVVL 960
QY 961 SVYDYMKAIGLFIPLSTFLPMCHNVHVSALASNYLSLWTDPIVNGTQEHKVKRLSVYG 1020
Db 961 SVYDYMKAIGLFIPLSTFLPMCHNVHVSALASNYLSLWTDPIVNGTQEHKVKRLSVYG 1020
QY 1021 ALGISQGIYAVFGYSMAVSGITGILASRCLHVDLLHSILRSPMSFFERTPSGMLVNRFSKEL 1080
Db 1021 ALGISQGIYAVFGYSMAVSGITGILASRCLHVDLLHSILRSPMSFFERTPSGMLVNRFSKEL 1080
QY 1081 DTVDSMIPEVIMKMGSLFNVTGACIVILLATPIAAIIIPPLGLIYFFVQRYVASSRQL 1140
Db 1081 DTVDSMIPEVIMKMGSLFNVTGACIVILLATPIAAIIIPPLGLIYFFVQRYVASSRQL 1140
QY 1141 KELESVSRSPVSHFNETLLGVSIVRAPEEQERFIHQSDLKVDENOKAYYPSIVANRWLA 1200
Db 1141 KELESVSRSPVSHFNETLLGVSIVRAPEEQERFIHQSDLKVDENOKAYYPSIVANRWLA 1200
QY 1201 VRLECVGNCIVLFAALFAVISRHSLSAGLVLSVSYSLQVTTYLNWLVMSSEMETNIVA 1260
Db 1201 VRLECVGNCIVLFAALFAVISRHSLSAGLVLSVSYSLQVTTYLNWLVMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKAPWQIOETAPPSPQVGRVFNRYCLRYREDLDFVLRHINVTINGG 1320
Db 1261 VERLKEYSETEKAPWQIOETAPPSPQVGRVFNRYCLRYREDLDFVLRHINVTINGG 1320
QY 1321 EKVGIVGRTGAGKSLTGLPRINESAGEIIIDGINIAKIGLHDLRKTIIIPQDPVLF 1380
Db 1321 EKVGIVGRTGAGKSLTGLPRINESAGEIIIDGINIAKIGLHDLRKTIIIPQDPVLF 1380
QY 1381 SGSLRWNLDPFSQYSDDEEVTWTSLELAHLKDFVSALPDKLDHCEAGGENLSVGQRQLVCL 1440
Db 1381 SGSLRWNLDPFSQYSDDEEVTWTSLELAHLKDFVSALPDKLDHCEAGGENLSVGQRQLVCL 1440
QY 1441 APALLRKTILVLDATAAVDLETDLLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500
Db 1441 APALLRKTILVLDATAAVDLETDLLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500
QY 1501 DKGEIOEQGAPSDLLQORGLFYSMADAGLV 1531
Db 1501 DKGEIOEQGAPSDLLQORGLFYSMADAGLV 1531

```

```

RESULT 10
US-08-460-907B-2
; Sequence 2, Application US/08460907B
; Patent No. 5891724
; GENERAL INFORMATION:
; APPLICANT: Deeley, Roger G.
; APPLICANT: Cole, Susan P.C.
; TITLE OF INVENTION: METHODS FOR CONFERRING MULTIDRUG
; TITLE OF INVENTION: RESISTANCE ON A CELL
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSES: PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS
; STREET: Queen's University at Kingston
; CITY: Kingston
; STATE: Ontario
; COUNTRY: CANADA

```

```

; ZIP: K7L 3N6
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,907B
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/966,923
; FILING DATE: 27-OCT-1992
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/029,340
; FILING DATE: 8-MAR-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/407,207
; FILING DATE: 20-MAR-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Steeg, Carol Miernicki
; REGISTRATION NUMBER: 39,539
; REFERENCE/DOCKET NUMBER: Q1551
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (613) 545-2342
; TELEFAX: (613) 545-6853
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1531 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-460-907B-2

```

```

Query Match 78.4%; Score 7849; DB 2; Length 1531;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MALRGFCSADGSDPLMDWNTWNTSNPDFTKCFQNTLVVWVPCFYLWACFPFYLYLSRH 60
Db 1 MALRGFCSADGSDPLMDWNTWNTSNPDFTKCFQNTLVVWVPCFYLWACFPFYLYLSRH 60
QY 61 DRGYIQTPLNKTALGFLWIVCWADLFYSFWERSRGIFLAPVFLVSPITLLGTTLLA 120
Db 61 DRGYIQTPLNKTALGFLWIVCWADLFYSFWERSRGIFLAPVFLVSPITLLGTTLLA 120
QY 121 TFLQLERRKGVQSSGIMLTFWLVALVCALLAILRSKINTALKEDAQVDLFRDITFYVYFS 180
Db 121 TFLQLERRKGVQSSGIMLTFWLVALVCALLAILRSKINTALKEDAQVDLFRDITFYVYFS 180
QY 181 LLLQLVLSCFSDRSPLFSETHDHPNCPSSASFLSRITFWIITGLIVRGVQPLEGSD 240
Db 181 LLLQLVLSCFSDRSPLFSETHDHPNCPSSASFLSRITFWIITGLIVRGVQPLEGSD 240
QY 241 LWSLNKEDTSBQVVPVLVKNMKCEAKTRKQPKVYVSSKDPAPKSSKVDANEVEAL 300
Db 241 LWSLNKEDTSBQVVPVLVKNMKCEAKTRKQPKVYVSSKDPAPKSSKVDANEVEAL 300
QY 301 IVKSPQKEWNPFLFKVLYKTPGIFLMSFFPKAIDLMMFSGPOLKLLIKFVNDTKAPD 360
Db 301 IVKSPQKEWNPFLFKVLYKTPGIFLMSFFPKAIDLMMFSGPOLKLLIKFVNDTKAPD 360
QY 361 WQGYFTVLLFVTACLOTLVHOFHICFVSGMRIKTAIVIGAVYRKALVITNSARKSSTV 420
Db 361 WQGYFTVLLFVTACLOTLVHOFHICFVSGMRIKTAIVIGAVYRKALVITNSARKSSTV 420

```

QY 421 GEIVNLSVDAORFMDLATYINWISAPLOVILALYLLWNLGSPVLAVVAVVWLVMPVN 480  
 DB 421 GEIVNLSVDAORFMDLATYINWISAPLOVILALYLLWNLGSPVLAVVAVVWLVMPVN 480  
 QY 481 AVAMKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKLYAWELAFKDKVLAIQEBELVKL 540  
 DB 481 AVAMKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKLYAWELAFKDKVLAIQEBELVKL 540  
 QY 541 KSNYLSAVCTFTWCTPFLVALCTFAVYTTIDENNILDAQTAFVSLALFNILFPLNLI 600  
 DB 541 KSNYLSAVCTFTWCTPFLVALCTFAVYTTIDENNILDAQTAFVSLALFNILFPLNLI 600  
 QY 601 MWISSIVQASVSLKRLRIFLSHEELPDSIERPVKGGTNSITVRNATFTWASDPT 660  
 DB 601 MWISSIVQASVSLKRLRIFLSHEELPDSIERPVKGGTNSITVRNATFTWASDPT 660  
 QY 661 LINGITSIPEGALVAVVGVQCGKSLLSALLAEMDKVEGHVAIKGSVAVVPQAWIQND 720  
 DB 661 LINGITSIPEGALVAVVGVQCGKSLLSALLAEMDKVEGHVAIKGSVAVVPQAWIQND 720  
 QY 721 SURENLLFCQLEEPYRVSIOACALLPDLLEILPSGDRTEIGEKGVLNLSGGOKQVSLAR 780  
 DB 721 SURENLLFCQLEEPYRVSIOACALLPDLLEILPSGDRTEIGEKGVLNLSGGOKQVSLAR 780  
 QY 781 AVYSNADIYLFDDPLSAVDAHVGHKHFENVIGPKGMLKNKTRILVTHSMSYLPQVDVIV 840  
 DB 781 AVYSNADIYLFDDPLSAVDAHVGHKHFENVIGPKGMLKNKTRILVTHSMSYLPQVDVIV 840  
 QY 841 MSGGKISEMGVSOELLARDGAFAEFLRTYASTEQODAEENGVTGVSFGKGAQWENG 900  
 DB 841 MSGGKISEMGVSOELLARDGAFAEFLRTYASTEQODAEENGVTGVSFGKGAQWENG 900  
 QY 901 LVTDSAGKLOLQOLSSSSYSYSDISRHHNSTAELOKAEAKKETWKLMEADKAQGVKL 960  
 DB 901 LVTDSAGKLOLQOLSSSSYSYSDISRHHNSTAELOKAEAKKETWKLMEADKAQGVKL 960  
 QY 961 SVYWDYMKAIGLFISFLSIFLPMCNHVSALAGNYLWLTDDPIVNGTQEHKVKVLSVG 1020  
 DB 961 SVYWDYMKAIGLFISFLSIFLPMCNHVSALAGNYLWLTDDPIVNGTQEHKVKVLSVG 1020  
 QY 1021 ALGISQGIATVFGYSMAVSTGGILASRCLHVDLLHLSLRSPMFFERTPSGNLVNRFSEL 1080  
 DB 1021 ALGISQGIATVFGYSMAVSTGGILASRCLHVDLLHLSLRSPMFFERTPSGNLVNRFSEL 1080  
 QY 1081 DTVDMSIPEVIMKFMGSLFNVIACIVILLATPIAAIIIPPLGLIYFFVQRYVASSRQL 1140  
 DB 1081 DTVDMSIPEVIMKFMGSLFNVIACIVILLATPIAAIIIPPLGLIYFFVQRYVASSRQL 1140  
 QY 1141 KXLESVRSVPVSHFNETLLGVSVIRAFEEQERFIHQSDLVKVDENQKAYYPSIVANRWLA 1200  
 DB 1141 KXLESVRSVPVSHFNETLLGVSVIRAFEEQERFIHQSDLVKVDENQKAYYPSIVANRWLA 1200  
 QY 1201 VRLKXSETEKEAPWQIETAPPSSWPQGVGRVFNRYCLRYREDLDVLRHINVTINGG 1260  
 DB 1201 VRLKXSETEKEAPWQIETAPPSSWPQGVGRVFNRYCLRYREDLDVLRHINVTINGG 1260  
 QY 1261 VERLKEYSETEKEAPWQIETAPPSSWPQGVGRVFNRYCLRYREDLDVLRHINVTINGG 1320  
 DB 1261 VERLKEYSETEKEAPWQIETAPPSSWPQGVGRVFNRYCLRYREDLDVLRHINVTINGG 1320  
 QY 1321 EKVIGVGRGACKSLTGLFRINESAGEEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
 DB 1321 EKVIGVGRGACKSLTGLFRINESAGEEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
 QY 1381 SGLSRNLDPPSOYSDDEWTSLELAHLKDFVSALPKDLHECARGGENLSVGORQLVCL 1440  
 DB 1381 SGLSRNLDPPSOYSDDEWTSLELAHLKDFVSALPKDLHECARGGENLSVGORQLVCL 1440  
 QY 1441 APALLRKTILVLDATAVLETDLLIQSTIRTOFEDCTVLTIAHRLNTMDYTRIVL 1500  
 DB 1441 APALLRKTILVLDATAVLETDLLIQSTIRTOFEDCTVLTIAHRLNTMDYTRIVL 1500  
 QY 1501 DKGEIQEYCAPSDLLIQORGLFYSMKADAGLV 1531

DB 1501 DKGEIQEYCAPSDLLIQORGLFYSMKADAGLV 1531  
 RESULT 11  
 US-08-463-179A-2  
 ; Sequence 2, Application US/08463179A  
 ; Patent No. 6001563  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cole, Susan P.C.  
 ; APPLICANT: Deeley, Roger G.  
 ; TITLE OF INVENTION: METHODS FOR IDENTIFYING CHEMOSENSITIZERS  
 ; NUMBER OF SEQUENCES: 6  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESS: LAHIVE & COCKFIELD  
 ; STREET: 60 State Street, suite 510  
 ; CITY: Boston  
 ; STATE: Massachusetts  
 ; COUNTRY: USA  
 ; ZIP: 02109  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: ASCII text  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/463,179A  
 ; FILING DATE:  
 ; CLASSIFICATION: 536  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 07/966,923  
 ; FILING DATE: 27-OCT-1992  
 ; APPLICATION NUMBER: 08/029,340  
 ; FILING DATE: 8-MAR-1993  
 ; APPLICATION NUMBER: 08/141,893  
 ; FILING DATE: 26-OCT-1993  
 ; APPLICATION NUMBER: 08/407,207  
 ; FILING DATE: 20-MAR-1995  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: DeConti, Giulio A. Jr.  
 ; REGISTRATION NUMBER: 31,503  
 ; REFERENCE/DOCKET NUMBER: PQ1-002CP8  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (617) 227-7400  
 ; TELEFAX: (617) 227-5941  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1531 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-463-179A-2  
 Query Match 78.4%; Score 7849; DB 3; Length 1531;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 MALRGFCADGSDPLMDWNTWNTSNPDFTKCFONTVLVWVPCFYLWACFPFFFLYLSRH 60  
 DB 1 MALRGFCADGSDPLMDWNTWNTSNPDFTKCFONTVLVWVPCFYLWACFPFFFLYLSRH 60  
 QY 61 DRGIQMTPLNKTALGFLLLWIVCMADLFVSPWERSRGIFLAPVFLVSPFLGTTLLA 120  
 DB 61 DRGIQMTPLNKTALGFLLLWIVCMADLFVSPWERSRGIFLAPVFLVSPFLGTTLLA 120  
 QY 121 TFLQLERRKGVQSSGIMLTFWLVALCALAILRSKIMTALKEDAQVDLFRDITFYVYFS 180  
 DB 121 TFLQLERRKGVQSSGIMLTFWLVALCALAILRSKIMTALKEDAQVDLFRDITFYVYFS 180  
 QY 181 LLLIQLVLSCFSDRSPFLSETIHPNCPRESSASFLSRITFWWITGLIVRGYRQPLEGSD 240  
 DB 181 LLLIQLVLSCFSDRSPFLSETIHPNCPRESSASFLSRITFWWITGLIVRGYRQPLEGSD 240

QY 241 LMSLNKEDTSEQVVPVLVKNWKECAKTRKQPVVYYSKODPAQPKESKVDANEVEAL 300  
DB 241 LMSLNKEDTSEQVVPVLVKNWKECAKTRKQPVVYYSKODPAQPKESKVDANEVEAL 300  
QY 301 IVKSPKQEWNPGLFKVLKXTGPGPYFLMSPPFKAIHDLMMFSGPQILKLLIKFVNDTKAPD 360  
DB 301 IVKSPKQEWNPGLFKVLKXTGPGPYFLMSPPFKAIHDLMMFSGPQILKLLIKFVNDTKAPD 360  
QY 361 WQGYFYTVLLFVTFACLOTLVHLQYFHCIFVSGMRKTAIVIGAVYRKALVITNSARKSSTV 420  
DB 361 WQGYFYTVLLFVTFACLOTLVHLQYFHCIFVSGMRKTAIVIGAVYRKALVITNSARKSSTV 420  
QY 421 GEIVNLSVDQAFMDLATYINNIWSAPLOVLALYLWNLNLPVSLAGVAVVWLMVFN 480  
DB 421 GEIVNLSVDQAFMDLATYINNIWSAPLOVLALYLWNLNLPVSLAGVAVVWLMVFN 480  
QY 481 AVWAKTKTYQVAHMKSKDNRIKLNNEILNGIKVLKYAWELA FKDKVLAIROELKVLK 540  
DB 481 AVWAKTKTYQVAHMKSKDNRIKLNNEILNGIKVLKYAWELA FKDKVLAIROELKVLK 540  
QY 541 KSAYLSAVGTFTWCTPFLVALCTPAVYVTTIDENNILDAQTAFAVSLALFNILRFPNLILP 600  
DB 541 KSAYLSAVGTFTWCTPFLVALCTPAVYVTTIDENNILDAQTAFAVSLALFNILRFPNLILP 600  
QY 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKOGGTTNSITVRNATFTWASDPPT 660  
DB 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKOGGTTNSITVRNATFTWASDPPT 660  
QY 661 LINGITFISPEGALVAVGVGCGKSLLSALLAEMDKVEGHVAIKGSVAVYVQQAQWIND 720  
DB 661 LINGITFISPEGALVAVGVGCGKSLLSALLAEMDKVEGHVAIKGSVAVYVQQAQWIND 720  
QY 721 SURENILFCOLEEPPYRSVIOACALLPDLLEILPSGDRTEIGEKGWNLSSGGQKQVSLAR 780  
DB 721 SURENILFCOLEEPPYRSVIOACALLPDLLEILPSGDRTEIGEKGWNLSSGGQKQVSLAR 780  
QY 781 AVYSNADIYLPDPLSADVAHVGKHFENVIGPKGMLKNKTRILVTHSMSYLPQVDVIV 840  
DB 781 AVYSNADIYLPDPLSADVAHVGKHFENVIGPKGMLKNKTRILVTHSMSYLPQVDVIV 840  
QY 841 MSGGKISEMGSYQELLARDGAFELRTYASTEQDABENGVTGVSFGPKGKAKOMENGM 900  
DB 841 MSGGKISEMGSYQELLARDGAFELRTYASTEQDABENGVTGVSFGPKGKAKOMENGM 900  
QY 901 LVTDSAGKOLQOLQSSSSSYSGDISRHNSSTAELKAEAKETWKLMEADKAQOGVKL 960  
DB 901 LVTDSAGKOLQOLQSSSSSYSGDISRHNSSTAELKAEAKETWKLMEADKAQOGVKL 960  
QY 961 SVYDYMKAIGLFI8FLSIFLPMCNHVSALASNYWLSLWTDPIVNGTQEHKVRLSVYG 1020  
DB 961 SVYDYMKAIGLFI8FLSIFLPMCNHVSALASNYWLSLWTDPIVNGTQEHKVRLSVYG 1020  
QY 1021 ALGISQGIATVFGYSMAVSI8GILASCLHVDLHLSILSPMSFPFRTSGNIVNRFPSKEL 1080  
DB 1021 ALGISQGIATVFGYSMAVSI8GILASCLHVDLHLSILSPMSFPFRTSGNIVNRFPSKEL 1080  
QY 1081 DTVDMSIPVIMKFMGSLFNVIACIVILLATPIAAIIPPLGLIYFFVQRFYVASSRQL 1140  
DB 1081 DTVDMSIPVIMKFMGSLFNVIACIVILLATPIAAIIPPLGLIYFFVQRFYVASSRQL 1140  
QY 1141 KRLESVSRSPVYSHNETLLGVSVIRAFEEQERFIHQSDLKVDENQKAYPSIVANRWLA 1200  
DB 1141 KRLESVSRSPVYSHNETLLGVSVIRAFEEQERFIHQSDLKVDENQKAYPSIVANRWLA 1200  
QY 1261 VRLKEYSTEKAPMIOETAPPSSWPQGVGFERNYCLRYREDLDFVLRHINTINGG 1320  
DB 1261 VRLKEYSTEKAPMIOETAPPSSWPQGVGFERNYCLRYREDLDFVLRHINTINGG 1320  
QY 1321 EKVGI VGR TGAGKSSITLGLFRINESAGEI IIDGINIAKIGLHDLRFKTIIPQDPVLF 1380

DB 1321 EKVGI VGR TGAGKSSITLGLFRINESAGEI IIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
QY 1381 SGLSRMNLDPFSQYSDDEEVWTSLELAHLKDFVSALPDKLDHECAGGENLSVGQQLVCL 1440  
DB 1381 SGLSRMNLDPFSQYSDDEEVWTSLELAHLKDFVSALPDKLDHECAGGENLSVGQQLVCL 1440  
QY 1441 ARALLRKTILVLEDEATAAVALDLEDDLIIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIL 1500  
DB 1441 ARALLRKTILVLEDEATAAVALDLEDDLIIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIL 1500  
QY 1501 DKGEIQEYGAPSDILLOQRGLFYSMAKDAGLV 1531  
DB 1501 DKGEIQEYGAPSDILLOQRGLFYSMAKDAGLV 1531  
RESULT 12  
US-08-461-384B-2  
; Sequence 2, Application US/08461384B  
; Patent No. 6025473  
; GENERAL INFORMATION:  
; APPLICANT: Cole, Susan P.C.  
; APPLICANT: Deeley, Roger G.  
; TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS  
; STREET: Queen's University at Kingston  
; CITY: Kingston  
; STATE: Ontario  
; COUNTRY: CANADA  
; ZIP: K7L 3N6  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/461,384B  
; FILING DATE: 05-JUN-95  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/966,923  
; FILING DATE: 27-OCT-1992  
; APPLICATION NUMBER: 08/029,340  
; FILING DATE: 8-MAR-1993  
; APPLICATION NUMBER: 08/141,893  
; FILING DATE: 26-OCT-1993  
; APPLICATION NUMBER: 08/407,207  
; FILING DATE: 20-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Steeg, Carol Miernicki  
; REGISTRATION NUMBER: 39,539  
; REFERENCE/DOCKET NUMBER: Q1547  
; TELEPHONE: (613) 545-2342  
; TELEFAX: (613) 545-6853  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1531 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-461-384B-2

Query Match 78.4%; Score 7849; DB 3; Length 1531;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 MALRGFCADGSDPLDWNWNTWNTSNPDTKCFQNTVLVWVPCFYLWACFPFFYLSRH 60  
DB 1 MALRGFCADGSDPLDWNWNTWNTSNPDTKCFQNTVLVWVPCFYLWACFPFFYLSRH 60  
QY 61 DRGYIQMTPLNKTALGFLWVWADLFYSFWSERSGIFLAPVFLVSPITLLGTTLLA 120

Db 61 DRGYI QMTPLNKTALGFLFWVADLFYSPWERSRGI FLAPFLVSP TLLGTTLLA 120  
 QY 121 TELIQLERRKGVOSSGIMLTFWLVALVLCALAILRSKIMTALKEADAQVDLFRDITFVYVFS 180  
 Db 121 TELIQLERRKGVOSSGIMLTFWLVALVLCALAILRSKIMTALKEADAQVDLFRDITFVYVFS 180  
 QY 181 LLLIQLVLCFSDRSPLFSETIHDNPNCKEAKTRKQPKVYVSSKDPAPKESKVDANEVEAL 240  
 Db 181 LLLIQLVLCFSDRSPLFSETIHDNPNCKEAKTRKQPKVYVSSKDPAPKESKVDANEVEAL 240  
 QY 241 LMSLNKEDTSEQVVPVVLVKNKKECAKTRKQPKVYVSSKDPAPKESKVDANEVEAL 300  
 Db 241 LMSLNKEDTSEQVVPVVLVKNKKECAKTRKQPKVYVSSKDPAPKESKVDANEVEAL 300  
 QY 301 IVKSPQKWNPSLFLVLYKTGPFYFLMSPFFKAIHDLMMFSGPQILKLLIKFVNDTKAPD 360  
 Db 301 IVKSPQKWNPSLFLVLYKTGPFYFLMSPFFKAIHDLMMFSGPQILKLLIKFVNDTKAPD 360  
 QY 361 WQGYFYTVLLFTVACLQTLVLHQQYFHI CFVSGMIRIKTAVIGAVYKALVITNSARKSSTV 420  
 Db 361 WQGYFYTVLLFTVACLQTLVLHQQYFHI CFVSGMIRIKTAVIGAVYKALVITNSARKSSTV 420  
 QY 421 GEIVNLMSVDAQRFMDLATYINMIWSAPLQVILALYLWNLGSPVLAGVAVWVLVMPVN 480  
 Db 421 GEIVNLMSVDAQRFMDLATYINMIWSAPLQVILALYLWNLGSPVLAGVAVWVLVMPVN 480  
 QY 481 AYWMKTKTYQVAHMKSKDNRIKLMEILNGIKVLKYAWELAFKDKVLAIROBELKVLK 540  
 Db 481 AYWMKTKTYQVAHMKSKDNRIKLMEILNGIKVLKYAWELAFKDKVLAIROBELKVLK 540  
 QY 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTIDENILDAQTAFAVSLALFNILRFPNLILP 600  
 Db 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTIDENILDAQTAFAVSLALFNILRFPNLILP 600  
 QY 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGCGTNSITVRNATFTWASDDPPT 660  
 Db 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGCGTNSITVRNATFTWASDDPPT 660  
 QY 661 LNGITFSPGALVAVGVGCGKSLLSALLAENDKVEGHVAIKGSVAVYVPOQAWIQND 720  
 Db 661 LNGITFSPGALVAVGVGCGKSLLSALLAENDKVEGHVAIKGSVAVYVPOQAWIQND 720  
 QY 721 SLRENILFGCOLLEPYRYSVIOACALLPDLPLEIPSGDRTEIGEGKVNLSGGOKQVSLAR 780  
 Db 721 SLRENILFGCOLLEPYRYSVIOACALLPDLPLEIPSGDRTEIGEGKVNLSGGOKQVSLAR 780  
 QY 781 AVYSNADIYLFDDPLSAVDHVGKHI FENVIGPKGMLKNKTRILVTHSMSYLPQVDVILV 840  
 Db 781 AVYSNADIYLFDDPLSAVDHVGKHI FENVIGPKGMLKNKTRILVTHSMSYLPQVDVILV 840  
 QY 841 MSGGKISEMGSYOELLARDGAPAEFLRTYASTEQODAEENGVTGVSFGKEAKOMENGM 900  
 Db 841 MSGGKISEMGSYOELLARDGAPAEFLRTYASTEQODAEENGVTGVSFGKEAKOMENGM 900  
 QY 901 LVTDGAGQLOQLSSSSSYSGDISRHNSIAELQKAEAKKEETWKLMEADKAOTGOVKL 960  
 Db 901 LVTDGAGQLOQLSSSSSYSGDISRHNSIAELQKAEAKKEETWKLMEADKAOTGOVKL 960  
 QY 961 SYVDYMKAIGLFTISFLSIFLPMCHVVSALASNYWLSLWTDPIVNGTOEHKTVRLSVYG 1020  
 Db 961 SYVDYMKAIGLFTISFLSIFLPMCHVVSALASNYWLSLWTDPIVNGTOEHKTVRLSVYG 1020  
 QY 1021 ALGISQIAGVGYSAVSGIGGLASRCLHVDLLHSILRSPMSFFERTPSGNLVRNFSKEL 1080  
 Db 1021 ALGISQIAGVGYSAVSGIGGLASRCLHVDLLHSILRSPMSFFERTPSGNLVRNFSKEL 1080  
 QY 1081 DTVDMSIPEVIMKFMGSLFNIVIGACIVILLATPIAAIIIPPLGLIYFFVQRYFYVASSRQL 1140  
 Db 1081 DTVDMSIPEVIMKFMGSLFNIVIGACIVILLATPIAAIIIPPLGLIYFFVQRYFYVASSRQL 1140  
 QY 1141 KRLSVSRSPVYSHNETLLGVSVIRAFEEQERFIHQSDLKVDENQKAYYPSIVANRWLA 1200

Db 1141 KRLSVSRSPVYSHNETLLGVSVIRAFEEQERFIHQSDLKVDENQKAYYPSIVANRWLA 1200  
 QY 1201 VRLECVGNCIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTYLNLWLRMSSEMETNI VA 1260  
 Db 1201 VRLECVGNCIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTYLNLWLRMSSEMETNI VA 1260  
 QY 1261 VERLKEYSETKEAPWQIOETAPPSSWPQVGRVEFRNYCLARYEDLDVFLRHINTYINGG 1320  
 Db 1261 VERLKEYSETKEAPWQIOETAPPSSWPQVGRVEFRNYCLARYEDLDVFLRHINTYINGG 1320  
 QY 1321 EKVGI VGTGACKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
 Db 1321 EKVGI VGTGACKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
 QY 1381 SGLSRMNLDPFSQYSDEEVWTSLELAHLKDFVSALPDKLDHECAEGGENLSVGQRLVCL 1440  
 Db 1381 SGLSRMNLDPFSQYSDEEVWTSLELAHLKDFVSALPDKLDHECAEGGENLSVGQRLVCL 1440  
 QY 1441 ARALLRKTILVDEATAVDETTDLLOSTIRTOFEDCTVLTIAHRLNTIMDYTRVILV 1500  
 Db 1441 ARALLRKTILVDEATAVDETTDLLOSTIRTOFEDCTVLTIAHRLNTIMDYTRVILV 1500  
 QY 1501 DKGEIQEYCAPSDLLQQRGLFYSMADAGLV 1531  
 Db 1501 DKGEIQEYCAPSDLLQQRGLFYSMADAGLV 1531

RESULT 13

US-08-407-207A-2  
 ; Sequence 2, Application US/08407207A  
 ; Patent No. 6063621  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Deeley, Roger G.  
 ; APPLICANT: Cole, Susan P.C.  
 ; TITLE OF INVENTION: ANTIBODIES TO A MULTIDRUG RESISTANCE PROTEIN  
 ; NUMBER OF SEQUENCES: 7  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS  
 ; STREET: Queen's University at Kingston  
 ; CITY: Kingston  
 ; STATE: Ontario  
 ; COUNTRY: CANADA  
 ; ZIP: K7L 3N6  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: ASCII text  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/407,207A  
 ; FILING DATE: 20-MAR-1995  
 ; CLASSIFICATION: 424  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 07/966,923  
 ; FILING DATE: 27-OCT-1992  
 ; APPLICATION NUMBER: 08/029,340  
 ; FILING DATE: 8-MAR-1993  
 ; APPLICATION NUMBER: 08/141,893  
 ; FILING DATE: 26-OCT-1993  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Steeg, Carol Miernicki  
 ; REGISTRATION NUMBER: 39,539  
 ; REFERENCE/DOCKET NUMBER: Q1512  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (613) 545-2342  
 ; TELEFAX: (613) 545-6853  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1531 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 US-08-407-207A-2

Query Match		78.4%;	Score 7849;	DB 3;	Length 1531;	
Best Local Similarity		99.9%;	Pred. No. 0;			
Matches 1529;		Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;	
Qy	1	MALRGCSDGSDPLWNVWNTWNTSNPDTKCFONTLVVWVFCFVLMACFPFYFLYLSRH	60			
Db	1	MALRGCSDGSDPLWNVWNTWNTSNPDTKCFONTLVVWVFCFVLMACFPFYFLYLSRH	60			
Qy	61	DRGYQMTPLNKTALGELLWVCWADLFYFWERSRGIFLAPVPLVSPFTLLGTTLLA	120			
Db	61	DRGYQMTPLNKTALGELLWVCWADLFYFWERSRGIFLAPVPLVSPFTLLGTTLLA	120			
Qy	121	TFLLQLERRKGVSQSGIMLTFMLVALVLCALAILRSKIMTALKEDAQVDLFRDITFVYFVS	180			
Db	121	TFLLQLERRKGVSQSGIMLTFMLVALVLCALAILRSKIMTALKEDAQVDLFRDITFVYFVS	180			
Qy	181	LLLIQLVLSCFSDRSPFSETHDNPCESSASFLSRITFWITGLIVRGVRQPLEGSD	240			
Db	181	LLLIQLVLSCFSDRSPFSETHDNPCESSASFLSRITFWITGLIVRGVRQPLEGSD	240			
Qy	241	LWSLNKEDTSEQVVPVVLVKNWKECAKTRKQPVKVYVSSKDPAPKESKVDANEVEAL	300			
Db	241	LWSLNKEDTSEQVVPVVLVKNWKECAKTRKQPVKVYVSSKDPAPKESKVDANEVEAL	300			
Qy	301	IVKSPQKEWNPFLKVLKTFPGPYFLMSPFFKAIHDLMMFSGPQILKLLIKFVNDTKAPD	360			
Db	301	IVKSPQKEWNPFLKVLKTFPGPYFLMSPFFKAIHDLMMFSGPQILKLLIKFVNDTKAPD	360			
Qy	361	WQGYFYTVLLFTVACLOTILVHQYHHI CFVSGMRKTAIVIGAVYRKALVITNSARKSSTV	420			
Db	361	WQGYFYTVLLFTVACLOTILVHQYHHI CFVSGMRKTAIVIGAVYRKALVITNSARKSSTV	420			
Qy	421	GIIVNLSMDAORFMDLATYINMISAPLOVILALYLLWNLGSPSLAGVAVWVLMVPVN	480			
Db	421	GIIVNLSMDAORFMDLATYINMISAPLOVILALYLLWNLGSPSLAGVAVWVLMVPVN	480			
Qy	481	AVMAMTKTKYQVAHMKSDNRKIKLMEILNGIKLVKLAYAWELAFKDKVLAIQEBELKVLK	540			
Db	481	AVMAMTKTKYQVAHMKSDNRKIKLMEILNGIKLVKLAYAWELAFKDKVLAIQEBELKVLK	540			
Qy	541	KSAYLSAVGTFTWCTPFFVALCTFAVYVTDENNILDAQTAFVSLALFNILRFPNLIPL	600			
Db	541	KSAYLSAVGTFTWCTPFFVALCTFAVYVTDENNILDAQTAFVSLALFNILRFPNLIPL	600			
Qy	601	MVSSIVQASVSLKRLRI FLSSHEELEDPSIERPVKGGGTNSITVRNATFTWARSDDPT	660			
Db	601	MVSSIVQASVSLKRLRI FLSSHEELEDPSIERPVKGGGTNSITVRNATFTWARSDDPT	660			
Qy	661	LANGITFSIPEGALVAVVGQVGCGKSLLSALLAEMDKVEGHVAIKGSVAVVPQQAQI QND	720			
Db	661	LANGITFSIPEGALVAVVGQVGCGKSLLSALLAEMDKVEGHVAIKGSVAVVPQQAQI QND	720			
Qy	721	SURENILFCQLEEPYRVSIVIOACALLPDLLEILPSGDRTEIGEKGVLNLSGGQKQVSLAR	780			
Db	721	SURENILFCQLEEPYRVSIVIOACALLPDLLEILPSGDRTEIGEKGVLNLSGGQKQVSLAR	780			
Qy	781	AVYSNADIYLFDDPLSADVAHVGKHLFENVIGPKMLKNKTRILVTHSHMSYLPQVDVILV	840			
Db	781	AVYSNADIYLFDDPLSADVAHVGKHLFENVIGPKMLKNKTRILVTHSHMSYLPQVDVILV	840			
Qy	841	MSGGKISEMGYSQELLARDGAFAEFLRTYASTEQEQAEEENGVTGVSGPGKEAKQWENG	900			
Db	841	MSGGKISEMGYSQELLARDGAFAEFLRTYASTEQEQAEEENGVTGVSGPGKEAKQWENG	900			
Qy	901	LVTDSAGLQQLQSSSSSYSGDISRHNSSTAELOKAEAKBETWKLMEADKAQTQGVKL	960			
Db	901	LVTDSAGLQQLQSSSSSYSGDISRHNSSTAELOKAEAKBETWKLMEADKAQTQGVKL	960			
Qy	961	SVYWDYMKAGLGFISPLSIFLFCMCHVSALSNYMLSLWTDPIVNGTOEHTKVRLSVYG	1020			
Db	961	SVYWDYMKAGLGFISPLSIFLFCMCHVSALSNYMLSLWTDPIVNGTOEHTKVRLSVYG	1020			

RESULT 14

US-08-463-092B-6  
; Sequence 6, Application US/08463092B  
; Patent No. 5766880  
; GENERAL INFORMATION:  
; APPLICANT: Cole, Susan P.C.  
; APPLICANT: Delevy, Roger G.  
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ENCODING  
; TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS  
; STREET: Queen's University at Kingston  
; CITY: Kingston  
; STATE: Ontario  
; COUNTRY: CANADA  
; ZIP: K7L 3N6  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/463,092B  
; FILING DATE: 05-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/966,923  
; FILING DATE: 27-OCT-1992  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/029,340  
; FILING DATE: 8-MAR-1993  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/141,893

Qy	1021	ALGISQGIAGVFGYSMAVSIIGGILASRCCLHVDLLHLSILRSPMGFFERTPSGNLVNRFSKEL	1080
Db	1021	ALGISQGIAGVFGYSMAVSIIGGILASRCCLHVDLLHLSILRSPMGFFERTPSGNLVNRFSKEL	1080
Qy	1081	DTVDSMIEPVTKMFNGSLFNVTGACIVILLATPIAAIIIPPLGLIYFFVQRFVASSROL	1140
Db	1081	DTVDSMIEPVTKMFNGSLFNVTGACIVILLATPIAAIIIPPLGLIYFFVQRFVASSROL	1140
Qy	1141	KELESVSRSPPVYSHFNETLLGVSIVIRAFEEQERFIHQSDLKVDENQKAYYPSIVANRWLA	1200
Db	1141	KELESVSRSPPVYSHFNETLLGVSIVIRAFEEQERFIHQSDLKVDENQKAYYPSIVANRWLA	1200
Qy	1201	VRLECVGNCIVLFAALFAVISRHSLSAGLVLSVYSQVTTYLNLVMSSEMETNIVA	1260
Db	1201	VRLECVGNCIVLFAALFAVISRHSLSAGLVLSVYSQVTTYLNLVMSSEMETNIVA	1260
Qy	1261	VERLKEYSETKEAPWQIOETAPPSSWPQVGRVFNRYCLRVREDLDVLRHINVTINGG	1320
Db	1261	VERLKEYSETKEAPWQIOETAPPSSWPQVGRVFNRYCLRVREDLDVLRHINVTINGG	1320
Qy	1321	EKVGIIVGRTGAGKSSLTGLFRINSAEAGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF	1380
Db	1321	EKVGIIVGRTGAGKSSLTGLFRINSAEAGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF	1380
Qy	1381	SGSLMMNLDPFSQYSDEEVTSLSLAHLKDFVSALPDKLDHECAGGENLSVGQRLVCL	1440
Db	1381	SGSLMMNLDPFSQYSDEEVTSLSLAHLKDFVSALPDKLDHECAGGENLSVGQRLVCL	1440
Qy	1441	ARALARKTKILVLDATAVDDTLQSTIRTOFEDCTVLTIAHRLNTINDYTRVIVL	1500
Db	1441	ARALARKTKILVLDATAVDDTLQSTIRTOFEDCTVLTIAHRLNTINDYTRVIVL	1500
Qy	1501	DKGEIQEYCAPSDLLQQRGLFYSMADAGLV	1531
Db	1501	DKGEIQEYCAPSDLLQQRGLFYSMADAGLV	1531

<p> ; FILING DATE: 26-OCT-1993  ; CLASSIFICATION: 435  ; PRIOR APPLICATION DATA:  ; APPLICATION NUMBER: 08/407,207  ; FILING DATE: 20-MAR-1995  ; CLASSIFICATION: 435  ; ATTORNEY/AGENT INFORMATION:  ; NAME: Steeg, Carol Miernicki  ; REGISTRATION NUMBER: 39,539  ; REFERENCE/DOCKET NUMBER: Q1546  ; TELECOMMUNICATION INFORMATION:  ; TELEPHONE: (613) 545-2342  ; TELEFAX: (613) 545-6853  ; INFORMATION FOR SEQ ID NO: 6:  ; SEQUENCE CHARACTERISTICS:  ; LENGTH: 1528 amino acids  ; TYPE: amino acid  ; TOPOLOGY: linear  ; MOLECULE TYPE: protein  US-08-463-0928-6 </p>		<p> Query Match 69.9%; Score 7002.5; DB 1; Length 1528;  Best Local Similarity 88.0%; Pred. No. 0;  Matches 1349; Conservative 102; Mismatches 75; Indels 7; Gaps 5; </p>	
QY	1	MALRGFCSDGSDPLMDNMTWNTSNPDFTKCFQNTVLVWPCFYLVWACFPFYLRLSRH	60
DB	1	MALRSPCSADGSDPLMDNMTWNTSNPDFTKCFQNTVLVWPCFYLVWACFPFYLRLSRH	60
QY	61	DRGYQMTPLNTKTALGFLIMVCMADLFYSFWERSRGIFLAPVLPSTLLGTTLLA	120
DB	61	DRGYQMTPLNTKTALGFLIMVCMADLFYSFWERSRGIFLAPVLPSTLLGTTLLA	120
QY	121	TELIOLERRKGVOSSGIMLTFWLWALCALATLRKIMTALKEDAQVDLFRDITFYVFS	180
DB	121	TELIOLERRKGVOSSGIMLTFWLWALCALATLRKIMTALKEDAQVDLFRDITFYVFS	180
QY	181	LLLIQLVLSGSDRPLSETHDPNCPPESSASFLSRITFWITGLIVRGYRQPLEGSD	240
DB	181	LVVLQVLSCFSDCSPLETHDPNCPPESSASFLSRITFWITGLIVRGYRQPLEGSD	240
QY	241	LWLNKEDTSEOVVPLVQNMKECAKTRKQPVVYIS-SKDPAQPKSSSKVDANEEVEA	299
DB	241	LWLNKEDTSEOVVPLVQNMKECAKTRKQPVVYIS-SKDPAQPKSSSKVDANEEVEA	299
QY	300	LIVKSPKEMNPLFKVLYKTFPGPYFLMSFPFKAHDLWMPFSGPOILKLLIKFVNDTKAP	359
DB	301	LIVKSPKEMNPLFKVLYKTFPGPYFLMSFPFKAHDLWMPFSGPOILKLLIKFVNDTKAP	359
QY	360	DWQGYFYTVLLFVTTACQLTLVHOFYHFCFVSGMRITKAVIGAVYRKALVITNARKSST	419
DB	361	DWQGYFYTVLLFVTTACQLTLVHOFYHFCFVSGMRITKAVIGAVYRKALVITNARKSST	419
QY	420	VGEIVNLMSVDAQRFMDLATYINNIWSAPLOVILALYLLWNLGSPVLAVGVMLVMPV	479
DB	421	VGEIVNLMSVDAQRFMDLATYINNIWSAPLOVILALYLLWNLGSPVLAVGVMLVMPV	479
QY	480	NAVMAKTKTYQVAHMKSKDNRIKLWNLGKIVLVYAWELAPKDKVLAIROBELKVL	539
DB	481	NAVMAKTKTYQVAHMKSKDNRIKLWNLGKIVLVYAWELAPKDKVLAIROBELKVL	539
QY	540	KGSAYLSAVGFTTWCTPFLVALCTFAVYVITDENNILDQAFTAVSLALFNILFPFLN	599
DB	541	KGSAYLSAVGFTTWCTPFLVALCTFAVYVITDENNILDQAFTAVSLALFNILFPFLN	599
QY	600	PWLTSSIVQASVSLKRLIFLSHEELEDPSDTERPVDKGGTNSITVNNATFTWARSDDP	659
DB	601	PWLTSSIVQASVSLKRLIFLSHEELEDPSDTERPVDKGGTNSITVNNATFTWARSDDP	659
QY	660	TLNGITFSPGALVAVGVQVCGCKSSLLSALLAEMDKVEGHVAKGSVAVVPOQAWION	719
DB	660	TLNGITFSPGALVAVGVQVCGCKSSLLSALLAEMDKVEGHVAKGSVAVVPOQAWION	719

RESULT 15  
US-08-462-109A-6  
; Sequence 6, Application US/08462109A  
; Patent No. 5882875  
; GENERAL INFORMATION:  
; APPLICANT: Cole, Susan P.C.  
; APPLICANT: Deeley, Roger G.  
; TITLE OF INVENTION: METHODS FOR IDENTIFYING  
; TITLE OF INVENTION: MULTIDRUG RESISTANT TUMOR CELLS  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street, suite 510  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA

```

;
; ZIP: 02109
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/462,109A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/966,923
; FILING DATE: 27-OCT-1992
; APPLICATION NUMBER: 08/029,340
; FILING DATE: 8-MAR-1993
; APPLICATION NUMBER: 08/141,893
; FILING DATE: 26-OCT-1993
; APPLICATION NUMBER: 08/407,207
; FILING DATE: 20-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: DeConti, Giulio A. Jr.
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: PQ1-002CP4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1528 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
; US-08-462-109A-6

```

```

Query Match 69.9%; Score 7002.5; DB 2; Length 1528;
Best Local Similarity 88.0%; Pred. No. 0;
Matches 1349; Conservative 102; Mismatches 75; Indels 7; Gaps 5;

QY 1 MALRGCSDAGDPLDWDNVNTNSPDFTKCFQNTLVWVPCFYLWACFPFYLGRH 60
DB 1 MALRGCSDAGDPLDWDNVNTNSPDFTKCFQNTLVWVPCFYLWACFPFYLGRH 60
QY 61 DRGYQMTPLNKTALGELLWVWADLFYFWSRSGIFLAPVFLVSPITLLGTTLLA 120
DB 61 DRGYQMTPLNKTALGELLWVWADLFYFWSRSGIFLAPVFLVSPITLLGTTLLA 120
QY 121 TFLIQLERRKGVOSSGIMLTFMLVALVCAIILRSKIMTALKEDAQVDLFRDITFYVPS 180
DB 121 TFLIQLERRKGVOSSGIMLTFMLVALVCAIILRSKIMTALKEDAQVDLFRDITFYVPS 180
QY 181 LLLIQLVLSGFCSDRSPFSETTHDNPCESSASFLSRITFWITGLIYRGVQPLEGSD 240
DB 181 LLLIQLVLSGFCSDRSPFSETTHDNPCESSASFLSRITFWITGLIYRGVQPLEGSD 240
QY 241 LWSLNKEDTSEQVVPVLVNNWKECAKTRKQPVVYIS-SKDPAPQKSSKYDANEVEEA 299
DB 241 LWSLNKEDTSEQVVPVLVNNWKECAKTRKQPVVYIS-SKDPAPQKSSKYDANEVEEA 299
QY 300 LIVKSPQKWNPSLFKVLKTPFPLMSFPFKAHDLMMFSGPQLKLLIKFVNDTKAP 359
DB 300 LIVKSPQKWNPSLFKVLKTPFPLMSFPFKAHDLMMFSGPQLKLLIKFVNDTKAP 359
QY 361 DMQGVFYTALLFVSACQLTALHQLYFHCIVSGMRIKTAVGAVYRKALVITNSAKSST 420
DB 361 DMQGVFYTALLFVSACQLTALHQLYFHCIVSGMRIKTAVGAVYRKALVITNSAKSST 420
QY 420 VGEIVNLSVDAQRFMDLATYINMWSAPLQVITLALYLWNLGSPVLAGVAVMLMVPV 479
DB 420 VGEIVNLSVDAQRFMDLATYINMWSAPLQVITLALYLWNLGSPVLAGVAVMLMVPV 479
QY 480 NAVAMKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYLAWELAFQDKVMSIQBELKVL 539
DB 480 NAVAMKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYLAWELAFQDKVMSIQBELKVL 540

```

```

QY 540 KKSAYLSAVGTFTWCTPFLVALCTFAVYVVTIDENNILDAQTAFVSLFALFNILRPLN 599
DB 540 KKSAYLSAVGTFTWCTPFLVALCTFAVYVVTIDENNILDAQTAFVSLFALFNILRPLN 599
QY 600 PMWISSIVQASVSLKRLRIFLSHEELEPDSIERPVKGGTNSITVNAFTWARSOPP 659
DB 600 PMWISSIVQASVSLKRLRIFLSHEELEPDSIERPVKGGTNSITVNAFTWARSOPP 659
QY 660 TNGITTSIPGALVAVVGVCGKSSLLSALLAEMDKVEGHVAIKGSVAVVYVQAWIQN 719
DB 660 TNGITTSIPGALVAVVGVCGKSSLLSALLAEMDKVEGHVAIKGSVAVVYVQAWIQN 719
QY 720 DSLRENTILFGCOLBEPYRSVIOCALLPDLLEILPSGDRTEIGEKGVLGGQKQVSLA 779
DB 720 DSLRENTILFGCOLBEPYRSVIOCALLPDLLEILPSGDRTEIGEKGVLGGQKQVSLA 779
QY 780 RAVYSNADIYLFDDPLSAVDAGVGHIFENYVPGKMLKNTKTRILVTHSMSVLPQVDVII 839
DB 780 RAVYSNADIYLFDDPLSAVDAGVGHIFENYVPGKMLKNTKTRILVTHSMSVLPQVDVII 839
QY 840 VMGGKISEMGSYOELLARDGAFAEFLRTYASTEOEQAEEENGVTGVSQPGKEAKQWENG 899
DB 840 VMGGKISEMGSYOELLARDGAFAEFLRTYASTEOEQAEEENGVTGVSQPGKEAKQWENG 899
QY 900 MLVTDGKQLQRLSSSSSYSGDISRHHNSTAELOKAEKKEETWKLMEADKAGTQVYK 959
DB 900 MLVTDGKQLQRLSSSSSYSGDISRHHNSTAELOKAEKKEETWKLMEADKAGTQVYK 959
QY 960 LSVTWYDKALGLFISFLSIFLPMCNHVSAALSNVLSLWTD--PIVNGTQEHKTVRLSV 1018
DB 960 LSVTWYDKALGLFISFLSIFLPMCNHVSAALSNVLSLWTD--PIVNGTQEHKTVRLSV 1018
QY 1019 YGALGISOGIAVFGYSMAVISGIGILLASRLCLHVDLHLSILRGPMSPFERTPSGNLVNRSK 1078
DB 1019 YGALGISOGIAVFGYSMAVISGIGILLASRLCLHVDLHLSILRGPMSPFERTPSGNLVNRSK 1078
QY 1079 ELDTVDSDMPEVIMKPMGSLFNIVGACIVILLATPIAAIIIPPLGLIYFFVQRFVASSR 1138
DB 1079 ELDTVDSDMPEVIMKPMGSLFNIVGACIVILLATPIAAIIIPPLGLIYFFVQRFVASSR 1138
QY 1139 QKRLSVRSRPVYSHFNETLLGVSVIRAFPEQERFIHQSDLKVDENKAYVPSIVANRW 1198
DB 1139 QKRLSVRSRPVYSHFNETLLGVSVIRAFPEQERFIHQSDLKVDENKAYVPSIVANRW 1198
QY 1199 LAVRLECVGNCIVLFAALPAVIRSHLSAGLVGLSVSYSLQITAYLNWLVNRSSEMETNI 1258
DB 1199 LAVRLECVGNCIVLFAALPAVIRSHLSAGLVGLSVSYSLQITAYLNWLVNRSSEMETNI 1258
QY 1259 VAVERLKEYSETEKAPWQIOETAPPSPQVGRVFRNYCLRYREDLDFVLRHINVTIN 1318
DB 1259 VAVERLKEYSETEKAPWQIOETAPPSPQVGRVFRNYCLRYREDLDFVLRHINVTIN 1318
QY 1319 GGEKGVIGRTGAGKSSITLGLFRINESAGEIIIDGINIAKIGLHDLRFXITIIPODVP 1378
DB 1319 GGEKGVIGRTGAGKSSITLGLFRINESAGEIIIDGINIAKIGLHDLRFXITIIPODVP 1378
QY 1379 LFSGSLRNWLPFSQYSDVEEYVMTSLEAHLKDFVSALPKLDHCEAEGENLSVGORQV 1438
DB 1379 LFSGSLRNWLPFSQYSDVEEYVMTSLEAHLKDFVSALPKLDHCEAEGENLSVGORQV 1438
QY 1439 CLARALLRKTILVLDEATAVADLETDLLIOSTIRTFQEDCTVLTIAHRLNTIMDYTRI 1498
DB 1439 CLARALLRKTILVLDEATAVADLETDLLIOSTIRTFQEDCTVLTIAHRLNTIMDYTRI 1498
QY 1499 VLDKGEIYQYGAPSDILLOOGLFYFYSMAKDAGLV 1531
DB 1499 VLDKGEIYQYGAPSDILLOOGLFYFYSMAKDAGLV 1531

```

Search completed: March 18, 2005, 11:03:47  
Job time : 48.7091 secs

**THIS PAGE BLANK (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2005, 11:03:55 ; Search time 615.662 Seconds  
(without alignments)  
1045.320 Million cell updates/sec

Title: US-10-665-283-8  
Perfect score: 10016  
Sequence: 1 MALRGFCSADGSDPLMDWNV.....RSVAVAKPKFSISPDLS 1947

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1401741 seqs, 330541175 residues

Total number of hits satisfying chosen parameters: 1401741

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : Published Applications AA:\*
- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
  - 2: /cgn2\_6/ptodata/2/pubpaa/ECT\_NEW\_PUB.pep.\*
  - 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
  - 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
  - 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
  - 6: /cgn2\_6/ptodata/2/pubpaa/ECTUS\_PUBCOMB.pep.\*
  - 7: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
  - 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
  - 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
  - 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
  - 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
  - 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
  - 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
  - 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
  - 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
  - 16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*
  - 17: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
  - 18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*
  - 19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
  - 20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	7860	78.5	1531	11	US-09-939-853A-86
2	7860	78.5	1531	16	US-10-667-891-6
3	7860	78.5	1531	17	US-10-484-577-678
4	7769	77.6	1515	11	US-09-939-853A-87
5	7037	70.3	1388	16	US-10-408-785A-1718
6	7002.5	69.9	1528	11	US-09-939-853A-88
7	4491.5	44.8	1527	11	US-09-939-853A-83
8	4491.5	44.8	1527	11	US-09-939-853A-84
9	4491.5	44.8	1527	15	US-10-295-027-1342
10	4396	43.9	1522	11	US-09-939-853A-85
11	3860.5	38.5	1548	16	US-10-667-891-1
12	3603.5	36.0	1505	16	US-10-667-891-4
13	3601.5	36.0	1499	16	US-10-667-891-3

14	3600	35.9	1542	15	US-10-363-112-4	Sequence 4, Appli
15	3600	35.9	1545	15	US-10-252-155-4	Sequence 4, Appli
16	3600	35.9	1545	15	US-10-252-155-24	Sequence 24, Appli
17	3600	35.9	1545	15	US-10-252-155-26	Sequence 26, Appli
18	3600	35.9	1545	15	US-10-252-155-28	Sequence 28, Appli
19	3600	35.9	1545	15	US-10-252-155-30	Sequence 30, Appli
20	3600	35.9	1545	15	US-10-252-155-32	Sequence 32, Appli
21	3600	35.9	1545	15	US-10-252-155-34	Sequence 34, Appli
22	3600	35.9	1545	15	US-10-252-155-38	Sequence 38, Appli
23	3600	35.9	1545	15	US-10-252-155-40	Sequence 40, Appli
24	3600	35.9	1545	15	US-10-252-155-46	Sequence 46, Appli
25	3600	35.9	1545	15	US-10-252-155-48	Sequence 48, Appli
26	3600	35.9	1545	15	US-10-363-112-2	Sequence 2, Appli
27	3600	35.9	1545	15	US-10-363-112-6	Sequence 6, Appli
28	3600	35.9	1545	15	US-10-363-112-8	Sequence 8, Appli
29	3600	35.9	1545	15	US-10-363-112-10	Sequence 10, Appli
30	3600	35.9	1545	15	US-10-363-112-12	Sequence 12, Appli
31	3600	35.9	1545	15	US-10-363-112-14	Sequence 14, Appli
32	3600	35.9	1545	15	US-10-363-112-16	Sequence 16, Appli
33	3599.5	35.9	1507	16	US-10-667-891-5	Sequence 5, Appli
34	3599	35.9	1545	15	US-10-252-155-42	Sequence 42, Appli
35	3597	35.9	1545	15	US-10-252-155-44	Sequence 44, Appli
36	3595	35.9	1545	16	US-10-741-601-369	Sequence 369, App
37	3595	35.9	1545	16	US-10-741-601-370	Sequence 370, App
38	3593	35.9	1545	15	US-10-252-155-36	Sequence 36, Appli
39	3582	35.8	1545	15	US-10-252-155-604	Sequence 604, App
40	3385.5	33.8	1503	10	US-09-792-616-3	Sequence 3, Appli
41	3385.5	33.8	1503	16	US-10-764-328-3	Sequence 3, Appli
42	3385.5	33.8	1503	17	US-10-741-600-1583	Sequence 1583, Ap
43	3289	32.8	1494	15	US-10-369-493-6697	Sequence 6697, Ap
44	3225	32.2	1573	15	US-10-369-493-6867	Sequence 6867, Ap
45	3225	32.2	1573	15	US-10-369-493-6868	Sequence 6868, Ap

ALIGNMENTS

RESULT 1  
US-09-939-853A-86  
; Sequence 86, Application US/09939853A  
; Publication No. US20040039163A1  
; GENERAL INFORMATION:  
; APPLICANT: Burgess et al.  
; TITLE OF INVENTION: No. US20040039163A1el Proteins and Nucleic Acids Encoding Same  
; FILE REFERENCE: 21402-099  
; CURRENT APPLICATION NUMBER: US/09/939,853A  
; CURRENT FILING DATE: 2001-08-27  
; PRIOR APPLICATION NUMBER: 60/228,191  
; PRIOR FILING DATE: 2000-08-25  
; PRIOR APPLICATION NUMBER: 60/267,300  
; PRIOR FILING DATE: 2001-02-08  
; PRIOR APPLICATION NUMBER: 60/269,961  
; PRIOR FILING DATE: 2001-02-20  
; PRIOR APPLICATION NUMBER: 60/277,337  
; PRIOR FILING DATE: 2001-03-20  
; NUMBER OF SEQ ID NOS: 159  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 86  
; LENGTH: 1531  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-939-853A-86

Query Match	78.5%	Score 7860;	DB 11;	Length 1531;
Best Local Similarity	100.0%;	Pred. No. 0;	Mismatches 0;	Indels 0;
Matches 1531;	Conservative	0;		Gaps 0;
Qy	1	MALRGFCSADGSDPLMDWNVTWNTSNPDFTKCFQNTLVWVPCFYLVWACFPFYLYLSH	60	
Db	1	MALRGFCSADGSDPLMDWNVTWNTSNPDFTKCFQNTLVWVPCFYLVWACFPFYLYLSH	60	
Qy	61	DRGIQMTPLNKTALGLFWLWVWADLFVSFWERSRGIFLWVPSPTLLGLTTLLA	120	

Db 61 DRGYIQWTPLNKTKTALGFLWVWADLPYSFWERSRGIFLAPVFLVPTLLGITLLA 120  
Qy 121 TFLIQLERRKGVOSSGIMLTFWLVALVCAIALIRSKIMTALKEDAQVDLFRDITFYVYFS 180  
Db 121 TFLIQLERRKGVOSSGIMLTFWLVALVCAIALIRSKIMTALKEDAQVDLFRDITFYVYFS 180  
Qy 181 LLLIQLVLSCFSDRSPLFSETIHDNPNCPRESSASFLSRITFWITGLIVRGYRQPLEGSD 240  
Db 181 LLLIQLVLSCFSDRSPLFSETIHDNPNCPRESSASFLSRITFWITGLIVRGYRQPLEGSD 240  
Qy 241 LWSLNKEDTSEQVVPVVLVKNWKECAKTRKQPVKVVYSKDPAPKPESSKVDANEVEAL 300  
Db 241 LWSLNKEDTSEQVVPVVLVKNWKECAKTRKQPVKVVYSKDPAPKPESSKVDANEVEAL 300  
Qy 301 IVKSPKQWNPFLFKVLYKTFGDFLMSFFFAIHLMMPSGPQIILKLIKFVNDTKAPD 360  
Db 301 IVKSPKQWNPFLFKVLYKTFGDFLMSFFFAIHLMMPSGPQIILKLIKFVNDTKAPD 360  
Qy 361 WQGYFTVLLFVTACIQTLVLHQYPHI CFVSGWRIKTA VIGAVYRKALVITNSARKSSTV 420  
Db 361 WQGYFTVLLFVTACIQTLVLHQYPHI CFVSGWRIKTA VIGAVYRKALVITNSARKSSTV 420  
Qy 421 GEIVNLSVDAQRFMDLATYINIMWSAPQVILALYLLNLGSPSLAGVAVMVLMPVN 480  
Db 421 GEIVNLSVDAQRFMDLATYINIMWSAPQVILALYLLNLGSPSLAGVAVMVLMPVN 480  
Qy 481 AVAMKTKTYQVAHMKSKDNRIKLANEILNGIKVLKYAWELAFKDKVLAIROBELKVLK 540  
Db 481 AVAMKTKTYQVAHMKSKDNRIKLANEILNGIKVLKYAWELAFKDKVLAIROBELKVLK 540  
Qy 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVITIDENNILDAQTAFVSLALFNILRFLNLP 600  
Db 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVITIDENNILDAQTAFVSLALFNILRFLNLP 600  
Qy 601 MVISSIVQASVSLKRLIRIFLSHEELEPDSIERPVPKOGGNTSIITVRNATFTWASDPT 660  
Db 601 MVISSIVQASVSLKRLIRIFLSHEELEPDSIERPVPKOGGNTSIITVRNATFTWASDPT 660  
Qy 661 LNGITFISPEGALVAVGVQVCGKSLLSALLAEMDKVEGHVAIKGSVAYVPOQAWIQND 720  
Db 661 LNGITFISPEGALVAVGVQVCGKSLLSALLAEMDKVEGHVAIKGSVAYVPOQAWIQND 720  
Qy 721 SURENIFGQCLEEPYRSVIOACALLPDLPLPSGDRTEIGEKGWNLGGOKQVSLAR 780  
Db 721 SURENIFGQCLEEPYRSVIOACALLPDLPLPSGDRTEIGEKGWNLGGOKQVSLAR 780  
Qy 781 AVYSNADIYLPDPLSADVAHVCKHI FENVIGPKGMLKNKTRILVTHSMYLPQVDVIV 840  
Db 781 AVYSNADIYLPDPLSADVAHVCKHI FENVIGPKGMLKNKTRILVTHSMYLPQVDVIV 840  
Qy 841 MSGGKISEMGSYOELLARDGAFELRTVASTEQDQAEENGVTGSGPGKEAKOMENG 900  
Db 841 MSGGKISEMGSYOELLARDGAFELRTVASTEQDQAEENGVTGSGPGKEAKOMENG 900  
Qy 901 LVTDAGKQLOQLSSSSSYSGDISRHNSSTAELQKAEAKKEETWKLMEADKAQTQVKL 960  
Db 901 LVTDAGKQLOQLSSSSSYSGDISRHNSSTAELQKAEAKKEETWKLMEADKAQTQVKL 960  
Qy 961 SVYDYMKAIGLFIISFLSIFLPMCNHVSALASNYWLSLWTDPIVNGQEHKVRLSYV 1020  
Db 961 SVYDYMKAIGLFIISFLSIFLPMCNHVSALASNYWLSLWTDPIVNGQEHKVRLSYV 1020  
Qy 1021 ALGISQGIAGVGSMAVSIIGTILASRCLHVLLHLSILRSPMFPERTPSGNLVNRFSEL 1080  
Db 1021 ALGISQGIAGVGSMAVSIIGTILASRCLHVLLHLSILRSPMFPERTPSGNLVNRFSEL 1080  
Qy 1081 DTVDNMIPEVIMKFMGSLFNIVIGACIVILLATPAAIIIPPLGLIYFFVQFYVASSRQL 1140  
Db 1081 DTVDNMIPEVIMKFMGSLFNIVIGACIVILLATPAAIIIPPLGLIYFFVQFYVASSRQL 1140  
Qy 1141 KRLESVSRPVSYPVSHNETLLGVSVIRAFEEQERFTHQSDLVKDNOKAYPSIVANRWLA 1200  
Db 1141 KRLESVSRPVSYPVSHNETLLGVSVIRAFEEQERFTHQSDLVKDNOKAYPSIVANRWLA 1200

Qy 1201 VRLCVCNVCIVLPAALFAVISHRSLSAGLVGLSVSVLSQVTTYLLNWLVRMSSEMETNIYA 1260  
Db 1201 VRLCVCNVCIVLPAALFAVISHRSLSAGLVGLSVSVLSQVTTYLLNWLVRMSSEMETNIYA 1260  
Qy 1261 VERLKEYSETEKEAPWQIQTETAPPSWPQVGRVEFRNYCLRYREDLDFVLRHINTYINGG 1320  
Db 1261 VERLKEYSETEKEAPWQIQTETAPPSWPQVGRVEFRNYCLRYREDLDFVLRHINTYINGG 1320  
Qy 1321 ERKGVIGRTGAGKSSILTLGLFRINESAGEIIIDGINIAKIGHDLRPFKTIIPDPPVLF 1380  
Db 1321 ERKGVIGRTGAGKSSILTLGLFRINESAGEIIIDGINIAKIGHDLRPFKTIIPDPPVLF 1380  
Qy 1381 SSSLRMNLDPPSQYSDEEVWTSLELAHLKDFVSALPDKLDHECAGGENLSVGQRLVCL 1440  
Db 1381 SSSLRMNLDPPSQYSDEEVWTSLELAHLKDFVSALPDKLDHECAGGENLSVGQRLVCL 1440  
Qy 1441 ARALLRKTKILVLDDEATAVDETDLLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVIL 1500  
Db 1441 ARALLRKTKILVLDDEATAVDETDLLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVIL 1500  
Qy 1501 DKGEIQEYCAPSDLLIQRGLFYSMADAGLV 1531  
Db 1501 DKGEIQEYCAPSDLLIQRGLFYSMADAGLV 1531  
RESULT 2  
US-10-667-891-6  
; Sequence 6, Application US/10667891  
; Publication No. US20040171024A1  
; GENERAL INFORMATION:  
; APPLICANT: ROTH, CHARLES W.  
; APPLICANT: BREY, PAUL T.  
; APPLICANT: HOLM, INGE  
; APPLICANT: GRAILLES, MARINE  
; APPLICANT: RZHETSKY, ANDREY  
; TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS IN DROSOPHILA AND  
; TITLE OF INVENTION: ANOPHELES  
; FILE REFERENCE: 03495-0294-00000  
; CURRENT APPLICATION NUMBER: US/10/667,891  
; CURRENT FILING DATE: 2003-09-23  
; PRIOR APPLICATION NUMBER: 60/413,469  
; PRIOR FILING DATE: 2002-09-26  
; NUMBER OF SEQ ID NOS: 76  
; SOFTWARE: Patent in Ver. 3.2  
; SEQ ID NO 6  
; LENGTH: 1531  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-667-891-6  
Query Match 78.5%; Score 7860; DB 16; Length 1531;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MALRGFCSADGSDPLDWNVTWNTSNPDFTKCFONTVLVWVPCFYLWACFPFYLYLSRH 60  
Db 1 MALRGFCSADGSDPLDWNVTWNTSNPDFTKCFONTVLVWVPCFYLWACFPFYLYLSRH 60  
Qy 61 DRGYIQWTPLNKTKTALGFLWVWADLPYSFWERSRGIFLAPVFLVSPPTLLGITLLA 120  
Db 61 DRGYIQWTPLNKTKTALGFLWVWADLPYSFWERSRGIFLAPVFLVSPPTLLGITLLA 120  
Qy 121 TFLIQLERRKGVOSSGIMLTFWLVALVCAIALIRSKIMTALKEDAQVDLFRDITFYVYFS 180  
Db 121 TFLIQLERRKGVOSSGIMLTFWLVALVCAIALIRSKIMTALKEDAQVDLFRDITFYVYFS 180  
Qy 181 LLLIQLVLSCFSDRSPLFSETIHDNPNCPRESSASFLSRITFWITGLIVRGYRQPLEGSD 240  
Db 181 LLLIQLVLSCFSDRSPLFSETIHDNPNCPRESSASFLSRITFWITGLIVRGYRQPLEGSD 240  
Qy 241 LWSLNKEDTSEQVVPVVLVKNWKECAKTRKQPVKVVYSKDPAPKPESSKVDANEVEAL 300  
Db 241 LWSLNKEDTSEQVVPVVLVKNWKECAKTRKQPVKVVYSKDPAPKPESSKVDANEVEAL 300

Db 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPVVYSSKDPAPQKESKVDANEVEAL 300  
 QY 301 IVKSPQKWNPSLKVLYKTFPGPYFLMSPFFKAIHDLMMFSGPQILKLLIKFVNDTKAPD 360  
 Db 301 IVKSPQKWNPSLKVLYKTFPGPYFLMSPFFKAIHDLMMFSGPQILKLLIKFVNDTKAPD 360  
 QY 361 WQGYFYTVLLFVTTACQLTLVHLHOFYHICFVSGMRKTAIVGAVYRKALVITNSARKSTV 420  
 Db 361 WQGYFYTVLLFVTTACQLTLVHLHOFYHICFVSGMRKTAIVGAVYRKALVITNSARKSTV 420  
 QY 421 GEIVNLSMSVDAORFMDLATYINMIWSAPLOVILALYLWNLGSPSLAGVAVMLMVPVN 480  
 Db 421 GEIVNLSMSVDAORFMDLATYINMIWSAPLOVILALYLWNLGSPSLAGVAVMLMVPVN 480  
 QY 481 AVAMKTKTYQVAHMKSKDNRIKLNEILNGIKVLKLYAWELAFKDKVLAIROBELKVLK 540  
 Db 481 AVAMKTKTYQVAHMKSKDNRIKLNEILNGIKVLKLYAWELAFKDKVLAIROBELKVLK 540  
 QY 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTTIDENNILDAQTAFAVSLALFNILRPFNLILP 600  
 Db 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTTIDENNILDAQTAFAVSLALFNILRPFNLILP 600  
 QY 601 MYISSIVQASVSLKRLIFLSHEELEPDSIERPVKGGGTSITVRNATFTWASDPT 660  
 Db 601 MYISSIVQASVSLKRLIFLSHEELEPDSIERPVKGGGTSITVRNATFTWASDPT 660  
 QY 661 LMGITFSIPEGALVAVVGQVCGKSLLSALLAEMDKVEGHVAIKGSVAVYVQQAWIOND 720  
 Db 661 LMGITFSIPEGALVAVVGQVCGKSLLSALLAEMDKVEGHVAIKGSVAVYVQQAWIOND 720  
 QY 721 SURENILFCQLEEPYRSVIOACALLPDLEILPSGDRTEIGEKGVLNLSGGQKQVSLAR 780  
 Db 721 SURENILFCQLEEPYRSVIOACALLPDLEILPSGDRTEIGEKGVLNLSGGQKQVSLAR 780  
 QY 781 AVYSNADIYLPDDPLSAVDHVGKHI FENVIGPKMLKNKTRILVTHSMSYLPQVDVILV 840  
 Db 781 AVYSNADIYLPDDPLSAVDHVGKHI FENVIGPKMLKNKTRILVTHSMSYLPQVDVILV 840  
 QY 841 MSGGKISEMGVQELLARCAFAELRTYASTEQEODAEENGVTGSGPKAKOMENGM 900  
 Db 841 MSGGKISEMGVQELLARCAFAELRTYASTEQEODAEENGVTGSGPKAKOMENGM 900  
 QY 901 LVTDSAGKQRLQSSSSSYSGDISRHNSSTAELQKAEAKKEETWKLMEADKAQTVQVKL 960  
 Db 901 LVTDSAGKQRLQSSSSSYSGDISRHNSSTAELQKAEAKKEETWKLMEADKAQTVQVKL 960  
 QY 961 SVYDWMKAIGLFI8PLSIFLFCMCHVSALASNYWLSLWTDPIVNGTQEHTKVRLSVYG 1020  
 Db 961 SVYDWMKAIGLFI8PLSIFLFCMCHVSALASNYWLSLWTDPIVNGTQEHTKVRLSVYG 1020  
 QY 1021 ALGISOGIAVFGYSMAVSI8GGLASRCLHVDLHLSILRSPMSFFERTPSGNLVNRFSEL 1080  
 Db 1021 ALGISOGIAVFGYSMAVSI8GGLASRCLHVDLHLSILRSPMSFFERTPSGNLVNRFSEL 1080  
 QY 1081 DTVDMSIMEVIMKMGSLFNVI8ACIVILLATPIAAIIPPLGLIYFFVQRFYVASSRQL 1140  
 Db 1081 DTVDMSIMEVIMKMGSLFNVI8ACIVILLATPIAAIIPPLGLIYFFVQRFYVASSRQL 1140  
 QY 1141 KRLESVSRSPVSHNETLLGVSIVTRAPEQBRFTHQSDLKVDENOKAYIPSIANRWLA 1200  
 Db 1141 KRLESVSRSPVSHNETLLGVSIVTRAPEQBRFTHQSDLKVDENOKAYIPSIANRWLA 1200  
 QY 1201 VRLECVGNCIVLFAALFAVISHRSLSAGLVGSVSYSLQVTTYLWNLVRMSMETNIVA 1260  
 Db 1201 VRLECVGNCIVLFAALFAVISHRSLSAGLVGSVSYSLQVTTYLWNLVRMSMETNIVA 1260  
 QY 1261 VERLKEYSETEAPWQI8ETAPPSSWPQVGRVEFRNYCLRYREDLDLVRHINVTINGG 1320  
 Db 1261 VERLKEYSETEAPWQI8ETAPPSSWPQVGRVEFRNYCLRYREDLDLVRHINVTINGG 1320  
 QY 1321 EKVGI8V8RTGAGKSSITLGLFRIN8SAGEI11DGINIAKIGLHDLRFKITIIPQDPVL 1380  
 Db 1321 EKVGI8V8RTGAGKSSITLGLFRIN8SAGEI11DGINIAKIGLHDLRFKITIIPQDPVL 1380

QY 1381 SCSLRMNLDPFSQXSDEEVWTSLELAHLKDFVSALPDKLDHECARGENLSYGQRLVCL 1440  
 Db 1381 SCSLRMNLDPFSQXSDEEVWTSLELAHLKDFVSALPDKLDHECARGENLSYGQRLVCL 1440  
 QY 1441 ARALLRKTKILVDEATAAANDLETTDDLIQSTIRTOFEDCTVLTIAHRLNTINDYTRVIL 1500  
 Db 1441 ARALLRKTKILVDEATAAANDLETTDDLIQSTIRTOFEDCTVLTIAHRLNTINDYTRVIL 1500  
 QY 1501 DKGEIOEGAPSDLLQORGLFYSMADAGLV 1531  
 Db 1501 DKGEIOEGAPSDLLQORGLFYSMADAGLV 1531

RESULT 3

US-10-484-577-678  
 ; Sequence 678, Application US/10484577  
 ; Publication No. US20050032724A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: EPIDAUROS Biotechnologie Aktiengesellschaft  
 ; TITLE OF INVENTION: Means and methods for improved treatment of cancer based on Ugr1  
 ; FILE REFERENCE: F2285PCT-1  
 ; CURRENT APPLICATION NUMBER: US/10/484,577  
 ; CURRENT FILING DATE: 2004-01-22  
 ; PRIOR APPLICATION NUMBER: PCT/EP 02/08220  
 ; PRIOR FILING DATE: 2002-07-23  
 ; PRIOR APPLICATION NUMBER: EP 01 11 7608.8  
 ; PRIOR FILING DATE: 2001-07-23  
 ; PRIOR APPLICATION NUMBER: EP 02011710.7  
 ; PRIOR FILING DATE: 2002-05-24  
 ; NUMBER OF SEQ ID NOS: 683  
 ; SOFTWARE: Patent in version 3.1  
 ; SEQ ID NO 678  
 ; LENGTH: 1531  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-484-577-678

Query Match 78.5%; Score 7860; DB 17; Length 1531;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCSDGSDPLWDNWNVTWNTSNPDTKCFQNTVLVWVPCFYLWACFPFYLYLSRH 60  
 Db 1 MALRGFCSDGSDPLWDNWNVTWNTSNPDTKCFQNTVLVWVPCFYLWACFPFYLYLSRH 60  
 QY 61 DRGYIOMTPLNKTALGFLWIVCWADIFYSFWERSRGIFLAPVFLVSPPTLLGTTLLA 120  
 Db 61 DRGYIOMTPLNKTALGFLWIVCWADIFYSFWERSRGIFLAPVFLVSPPTLLGTTLLA 120  
 QY 121 TPLIOLERRKGVQSSGIMLTFWLVALCALAILRSKIMTALKEDAQVDLFRDITFVYFS 180  
 Db 121 TPLIOLERRKGVQSSGIMLTFWLVALCALAILRSKIMTALKEDAQVDLFRDITFVYFS 180  
 QY 181 LLLIOLVLSGSDRSPFSETHDPNCPSSASFLSRITFWITGLIVRGYRQPLEGSD 240  
 Db 181 LLLIOLVLSGSDRSPFSETHDPNCPSSASFLSRITFWITGLIVRGYRQPLEGSD 240  
 QY 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPVVYSSKDPAPQKESKVDANEVEAL 300  
 Db 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPVVYSSKDPAPQKESKVDANEVEAL 300  
 QY 301 IVKSPQKWNPSLKVLYKTFPGPYFLMSPFFKAIHDLMMFSGPQILKLLIKFVNDTKAPD 360  
 Db 301 IVKSPQKWNPSLKVLYKTFPGPYFLMSPFFKAIHDLMMFSGPQILKLLIKFVNDTKAPD 360  
 QY 361 WQGYFYTVLLFVTTACQLTLVHLHOFYHICFVSGMRKTAIVGAVYRKALVITNSARKSTV 420  
 Db 361 WQGYFYTVLLFVTTACQLTLVHLHOFYHICFVSGMRKTAIVGAVYRKALVITNSARKSTV 420  
 QY 421 GEIVNLSMSVDAORFMDLATYINMIWSAPLOVILALYLWNLGSPSLAGVAVMLMVPVN 480  
 Db 421 GEIVNLSMSVDAORFMDLATYINMIWSAPLOVILALYLWNLGSPSLAGVAVMLMVPVN 480

481 AVAMKTKTYQVAHMKSKNRIKLMNEILNGIKVLKLYAWELAFKDKVLAIRQBELKVLK 540  
 481 AVAMKTKTYQVAHMKSKNRIKLMNEILNGIKVLKLYAWELAFKDKVLAIRQBELKVLK 540  
 541 KSAYLAVGFTTWCPTFLVALCTFAVYVITDENNILDQAQAFVSLALFNILRFFLNILP 600  
 541 KSAYLAVGFTTWCPTFLVALCTFAVYVITDENNILDQAQAFVSLALFNILRFFLNILP 600  
 601 MVLISSIVQASVSLKRLRIFLSHELEPDSIERRPVKGGGTNSITVRNATFTWASDPT 660  
 601 MVLISSIVQASVSLKRLRIFLSHELEPDSIERRPVKGGGTNSITVRNATFTWASDPT 660  
 661 LNCITPSIEGALVAVVGCGCKSSLLSALLAEMDKVGHVAIKGSVAVVQQAQWIND 720  
 661 LNCITPSIEGALVAVVGCGCKSSLLSALLAEMDKVGHVAIKGSVAVVQQAQWIND 720  
 721 SURENIFGQLEPYRYSVIOACALLPDLPLPSGDRTEIGEGKGNLSGGQKQVSLAR 780  
 721 SURENIFGQLEPYRYSVIOACALLPDLPLPSGDRTEIGEGKGNLSGGQKQVSLAR 780  
 781 AVYSNADIYLFDDPLSAVDHVKHIFENVIKPKMLKNKTRILVTHSMSYLPQVDVIV 840  
 781 AVYSNADIYLFDDPLSAVDHVKHIFENVIKPKMLKNKTRILVTHSMSYLPQVDVIV 840  
 841 MSGKISEMSGYQELIARDGAFAELRTYASTEQEOADAENGVTGSGPKGAQWENGM 900  
 841 MSGKISEMSGYQELIARDGAFAELRTYASTEQEOADAENGVTGSGPKGAQWENGM 900  
 901 LVTDSAGKQOLQSSSSSYSGDISRHNSSTAELOKAEAKKEETWKLMEADKAQGVKL 960  
 901 LVTDSAGKQOLQSSSSSYSGDISRHNSSTAELOKAEAKKEETWKLMEADKAQGVKL 960  
 961 SYVDYMKAGLFIPLSLFPLFMCNHSALASNYLSLWTDPIVNGTQEHKTVLSYV 1020  
 961 SYVDYMKAGLFIPLSLFPLFMCNHSALASNYLSLWTDPIVNGTQEHKTVLSYV 1020  
 1021 ALGISQIAGVFGYSMAVSIIGTILASRCLHVDLHLSILSPMSFFERTPSGNLVNRFSEL 1080  
 1021 ALGISQIAGVFGYSMAVSIIGTILASRCLHVDLHLSILSPMSFFERTPSGNLVNRFSEL 1080  
 1081 DTVDSMIEVVKMGFSLFNVTGACIVILLATPIAAIIIPGLIYFFVQRPYVASSRQL 1140  
 1081 DTVDSMIEVVKMGFSLFNVTGACIVILLATPIAAIIIPGLIYFFVQRPYVASSRQL 1140  
 1141 KELESVSRSPVYSHNETLLGVSVIRAFEEQERFIHQSDLVKDNOKAYPSIVANRWLA 1200  
 1141 KELESVSRSPVYSHNETLLGVSVIRAFEEQERFIHQSDLVKDNOKAYPSIVANRWLA 1200  
 1201 VRLECVGNCIVLFAALFAVISHRSLSAGLVGLSVSYSLQVTTYLNWLVNRMSEMETNIVA 1260  
 1201 VRLECVGNCIVLFAALFAVISHRSLSAGLVGLSVSYSLQVTTYLNWLVNRMSEMETNIVA 1260  
 1261 VERLKEYSETKEAPWQIQTAPPSSWPQGRVEFRNYCLIRYREDLDVPLRHINVTINGG 1320  
 1261 VERLKEYSETKEAPWQIQTAPPSSWPQGRVEFRNYCLIRYREDLDVPLRHINVTINGG 1320  
 1321 EKVGVGRTGAGKSLTLGLFRINESARGEIIIDGINIAKIGLHDLRPFKTIIPQDPVLF 1380  
 1321 EKVGVGRTGAGKSLTLGLFRINESARGEIIIDGINIAKIGLHDLRPFKTIIPQDPVLF 1380  
 1381 SGSLRMNLDPPSYQSDDEEVTSLAHLKDFVSAIPDKLDHECAGGENLSVGQRLVCL 1440  
 1381 SGSLRMNLDPPSYQSDDEEVTSLAHLKDFVSAIPDKLDHECAGGENLSVGQRLVCL 1440  
 1441 ARALLRKTILVLEATAVDLETDLLIQSTIRTPQEDCTVLTIAHRLNTIMDYTRIVL 1500  
 1441 ARALLRKTILVLEATAVDLETDLLIQSTIRTPQEDCTVLTIAHRLNTIMDYTRIVL 1500  
 1501 DKGEIQEYCAPSDLIQORGLFYSMAKDAGLV 1531  
 1501 DKGEIQEYCAPSDLIQORGLFYSMAKDAGLV 1531

RESULT 4  
 US-09-939-853A-87  
 ; Sequence 87, Application US/09939853A  
 ; Publication No. US20040039163A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Burgees et al.  
 ; TITLE OF INVENTION: No. US20040039163A1el Proteins and Nucleic Acids Encoding Same  
 ; FILE REFERENCE: 21402-099  
 ; CURRENT APPLICATION NUMBER: US/09/939,853A  
 ; CURRENT FILING DATE: 2001-08-27  
 ; PRIOR APPLICATION NUMBER: 60/228,191  
 ; PRIOR FILING DATE: 2000-08-25  
 ; PRIOR APPLICATION NUMBER: 60/267,300  
 ; PRIOR FILING DATE: 2001-02-08  
 ; PRIOR APPLICATION NUMBER: 60/269,961  
 ; PRIOR FILING DATE: 2001-02-20  
 ; PRIOR APPLICATION NUMBER: 60/277,337  
 ; PRIOR FILING DATE: 2001-03-20  
 ; NUMBER OF SEQ ID NOS: 159  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 87  
 ; TYPE: PRT  
 ; LENGTH: 1515  
 ; ORGANISM: Homo sapiens  
 ; US-09-939-853A-87

Query Match 77.6%; Score 7769; DB 11; Length 1515;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1515; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 DWNVTWNTSNDDFKCFQNTVVLVWVPCFYLWACFPFFLYLSRDRGVIQMTPLNKTXTA 76  
 DB 1 DWNVTWNTSNDDFKCFQNTVVLVWVPCFYLWACFPFFLYLSRDRGVIQMTPLNKTXTA 60

QY 77 LGFLLWIVCWADLFVSWERSRGIFLAPVFLVSPFLTGITTTLLATFLIQLERRGVQSSG 136  
 DB 61 LGFLLWIVCWADLFVSWERSRGIFLAPVFLVSPFLTGITTTLLATFLIQLERRGVQSSG 120

QY 137 IMLTFWLVALVCAALILRSKIMTALKEDAQVDLFRDITFVYVYFSLLLIQLVLSCFSDRSP 196  
 DB 121 IMLTFWLVALVCAALILRSKIMTALKEDAQVDLFRDITFVYVYFSLLLIQLVLSCFSDRSP 180

QY 197 LFSITIHDPNCPRESSASFSLRITFWITGLIVRGYQPLEGSDLSINKEDTSEQVVPV 256  
 DB 181 LFSITIHDPNCPRESSASFSLRITFWITGLIVRGYQPLEGSDLSINKEDTSEQVVPV 240

QY 257 LVKNWKECAKTRKQPVKVYVSSKDPAPKQESSKVDANEVEALIVKSPQKWNPSLPKV 316  
 DB 241 LVKNWKECAKTRKQPVKVYVSSKDPAPKQESSKVDANEVEALIVKSPQKWNPSLPKV 300

QY 317 LYKTFGYPFLMSPFFKAIHDLMMFSGPQILKLLIKFVNDTKAPDWQGYFYTVLLFVTACL 376  
 DB 301 LYKTFGYPFLMSPFFKAIHDLMMFSGPQILKLLIKFVNDTKAPDWQGYFYTVLLFVTACL 360

QY 377 QTLVLHOFYFHCFSVGMRIKTAVIGAVYRKALVITNSARKSSTVGEIVNLSVDAORFMD 436  
 DB 361 QTLVLHOFYFHCFSVGMRIKTAVIGAVYRKALVITNSARKSSTVGEIVNLSVDAORFMD 420

QY 437 LATYINMWSAPLQVTLALYLLMLNLGSPSLAGVAVMVLMPVNVNMAKTKTYQVAHMK 496  
 DB 421 LATYINMWSAPLQVTLALYLLMLNLGSPSLAGVAVMVLMPVNVNMAKTKTYQVAHMK 480

QY 497 SKDNRIKLMNEILNGIKVLKLYAWELAFKDKVLAIRQBELKVLKKSAYLSAVGFTTWCCT 556  
 DB 481 SKDNRIKLMNEILNGIKVLKLYAWELAFKDKVLAIRQBELKVLKKSAYLSAVGFTTWCCT 540

QY 557 PFLVALCTFAVYVITDENNILDQAQAFVSLALFNILRFFLNILPMWISSIVQASVSLKRL 616  
 DB 541 PFLVALCTFAVYVITDENNILDQAQAFVSLALFNILRFFLNILPMWISSIVQASVSLKRL 600

QY 617 RIFLSHELEBPDSIERRPVKGGGTNSITVRNATFTWASDPTLNGITFSIEGALVAV 676  
 DB 617 RIFLSHELEBPDSIERRPVKGGGTNSITVRNATFTWASDPTLNGITFSIEGALVAV 676

Db 601 RIFLSHEEPPDISERRPVKGGGNSITVRNATFTWASDPTLNGITFTSIPGALVAV 660  
Qy 677 VQVCGGKSSLLSALLAEMDKVEGHVAKGSVAYVPOQAWIQNDISRENILFGCOLEBPY 736  
Db 661 VQVCGGKSSLLSALLAEMDKVEGHVAKGSVAYVPOQAWIQNDISRENILFGCOLEBPY 720  
Qy 737 YRSVTOACALLPDLEILPSGDRTEIGEKNVLSGGQKORVSLARAVYSNADIIYLPDDPLS 796  
Db 721 YRSVTOACALLPDLEILPSGDRTEIGEKNVLSGGQKORVSLARAVYSNADIIYLPDDPLS 780  
Qy 797 AYDAHVGHKHFENVIGPKGMLKNKTRILVTHSMSYLPQVDVILVMSGGKISMSGYOELL 856  
Db 781 AYDAHVGHKHFENVIGPKGMLKNKTRILVTHSMSYLPQVDVILVMSGGKISMSGYOELL 840  
Qy 857 ARDGAFAEFLRYASTEOQDAEENGVTGSGPGKEAKOMENGLVTDTSAGKOLQRLSS 916  
Db 841 ARDGAFAEFLRYASTEOQDAEENGVTGSGPGKEAKOMENGLVTDTSAGKOLQRLSS 900  
Qy 917 SSSYSGDISRHNSHNSAELOKAEAKKEETWKLMEADKAQTQGVKLSVYWDYMKAIGLFISF 976  
Db 901 SSSYSGDISRHNSHNSAELOKAEAKKEETWKLMEADKAQTQGVKLSVYWDYMKAIGLFISF 960  
Qy 977 LSIFLPMCNHVSALASNYLSLWTDPIVNGTQHTKVLSTYVGALGISQGIATVFGSMA 1036  
Db 961 LSIFLPMCNHVSALASNYLSLWTDPIVNGTQHTKVLSTYVGALGISQGIATVFGSMA 1020  
Qy 1037 VSIGGILASRCLHVDLHLSILASPMSPFERTPSGNLVRFSKELDTVDSMIPEVIMPMWG 1096  
Db 1021 VSIGGILASRCLHVDLHLSILASPMSPFERTPSGNLVRFSKELDTVDSMIPEVIMPMWG 1080  
Qy 1097 SLFNVGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQLKRLSVSRSPYSHFN 1156  
Db 1081 SLFNVGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQLKRLSVSRSPYSHFN 1140  
Qy 1157 ETLGVSIVRAPEEQRPTHQSDLVKVDENOKAYPSIVANRWLAIVLSCVNCIVLFAAL 1216  
Db 1141 ETLGVSIVRAPEEQRPTHQSDLVKVDENOKAYPSIVANRWLAIVLSCVNCIVLFAAL 1200  
Qy 1217 FAVISRHLSLGLVLSYSIQVTTYLNWLRMSSEMETNIVAVERLKEYSETEKEAPW 1276  
Db 1201 FAVISRHLSLGLVLSYSIQVTTYLNWLRMSSEMETNIVAVERLKEYSETEKEAPW 1260  
Qy 1277 QIQTAPPSSWPQVGRVFRNCLAYREDLPVLRHINVTNGGKVGIVGRTGAGKSSL 1336  
Db 1261 QIQTAPPSSWPQVGRVFRNCLAYREDLPVLRHINVTNGGKVGIVGRTGAGKSSL 1320  
Qy 1337 TLGLFRINESAGEEIIIDGINIAKIGLHDLREKTIIPQDPVLFSGSLRMNLDPPSOYSD 1396  
Db 1321 TLGLFRINESAGEEIIIDGINIAKIGLHDLREKTIIPQDPVLFSGSLRMNLDPPSOYSD 1380  
Qy 1397 EEWMTSLELAHLKDFVSALPDKLDHECAEGGENLSVGORQLVCLARALLRKTILVLDEA 1456  
Db 1381 EEWMTSLELAHLKDFVSALPDKLDHECAEGGENLSVGORQLVCLARALLRKTILVLDEA 1440  
Qy 1457 TAAVDLETDLLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVILVDKGEIQEYGAPSDLLQ 1516  
Db 1441 TAAVDLETDLLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVILVDKGEIQEYGAPSDLLQ 1500  
Qy 1517 QRGLFYSMAKADAGLV 1531  
Db 1501 QRGLFYSMAKADAGLV 1515

## RESULT 5

US-10-408-765A-1718  
; Sequence 1718, Application US/10408765A  
; Publication No. US20040101874A1  
; GENERAL INFORMATION:  
; APPLICANT: Ghosh, Soumitra S.  
; APPLICANT: Fahy, Eoin D.  
; APPLICANT: Zhang, Bing  
; APPLICANT: Gibson, Bradford W.  
; APPLICANT: Taylor, Steven W.

; APPLICANT: Glenn, Gary M.  
; APPLICANT: Warnock, Dale E.  
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION  
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME  
; FILE REFERENCE: 660088.465  
; CURRENT APPLICATION NUMBER: US/10/408,765A  
; CURRENT FILING DATE: 2003-04-04  
; NUMBER OF SEQ ID NOS: 3077  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1718  
; LENGTH: 1388  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-408-765A-1718

Query Match 70.3%; Score 7037; DB 16; Length 1388;  
Best Local Similarity 97.0%; Pred. No. 0;

Matches 1387; Conservative 0; Mismatches 1; Indels 42; Gaps 1;

Qy 67 MTPLNKTALGFLLLWIVCWADLPYSPWERSRGIFLAPVFLVSPPTLLGITTLLATFLIQL 126  
Db 1 MTPLNKTALGFLLLWIVCWADLPYSPWERSRGIFLAPVFLVSPPTLLGITTLLATFLIQL 60  
Qy 127 ERRKGVQSSGIMLTFMLVALYCALAILRSKIMTALKEDAQVDLPDITFYVYFSLLLQL 186  
Db 61 ERRKGVQSSGIMLTFMLVALYCALAILRSKIMTALKEDAQVDLPDITFYVYFSLLLQL 97  
Qy 187 VLSCFSDRSPLFSETIHPNCPSSASFLSRITFWITGLIVRGYROPLEGSDLSLNK 246  
Db 98 -----NPCSSASFLSRITFWITGLIVRGYROPLEGSDLSLNK 138  
Qy 247 EDTSEQVVPVLVQNMKKECAKTRQPKVQVYSSKDPAPQKSSKVDANEEVEALIVKSPQ 306  
Db 139 EDTSEQVVPVLVQNMKKECAKTRQPKVQVYSSKDPAPQKSSKVDANEEVEALIVKSPQ 198  
Qy 307 KEWNPSPKVLKTFPGPYFLMSPPFKAIHDLMMSPGPOILKLLIKFVNDTKAPDQGGIFY 366  
Db 199 KEWNPSPKVLKTFPGPYFLMSPPFKAIHDLMMSPGPOILKLLIKFVNDTKAPDQGGIFY 258  
Qy 367 TVLLFVTAACLOTVLHQLVHFICFVSGMRIKTAIVGAVYRKALVITNSARKSSTVGEIVNL 426  
Db 259 TVLLFVTAACLOTVLHQLVHFICFVSGMRIKTAIVGAVYRKALVITNSARKSSTVGEIVNL 318  
Qy 427 MSVDAQRFMDLATYINMIWSAPLQVILALYLLNLGSPVLGAVVMLVMPVNAVMAK 486  
Db 319 MSVDAQRFMDLATYINMIWSAPLQVILALYLLNLGSPVLGAVVMLVMPVNAVMAK 378  
Qy 487 TKTVQVAHMKSDNRKILMNEILNGIKVLKYAMELAFKDKVLATROBELKVLKKSAYLS 546  
Db 379 TKTVQVAHMKSDNRKILMNEILNGIKVLKYAMELAFKDKVLATROBELKVLKKSAYLS 438  
Qy 547 AVGTFTVCTPFLVALCTFAVVTIDENNIIDAOATFVSLALFNILRPLNLPVWISSI 606  
Db 439 AVGTFTVCTPFLVALCTFAVVTIDENNIIDAOATFVSLALFNILRPLNLPVWISSI 498  
Qy 607 VQASVSLKRLRIFLSHEELEPDSIERREPVKDGGTNSITVRNATFTWASDPTLNGITF 666  
Db 499 VQASVSLKRLRIFLSHEELEPDSIERREPVKDGGTNSITVRNATFTWASDPTLNGITF 558  
Qy 667 SIPEGALVAVVGVQVCGKSSLLSALLAEMDKVEGHVAKGSVAYVPOQAWIQNDISREN 726  
Db 559 SIPEGALVAVVGVQVCGKSSLLSALLAEMDKVEGHVAKGSVAYVPOQAWIQNDISREN 618  
Qy 727 LFGCOLBEPYVRSVTOACALLPDLEILPSGDRTEIGEKNVLSGGQKORVSLARAVYNA 786  
Db 619 LFGCOLBEPYVRSVTOACALLPDLEILPSGDRTEIGEKNVLSGGQKORVSLARAVYNA 678  
Qy 787 DIYLFDDPLSAVDHVGKHFENVIGPKGMLKNKTRILVTHSMSYLPQVDVILVMSGGKI 846  
Db 679 DIYLFDDPLSAVDHVGKHFENVIGPKGMLKNKTRILVTHSMSYLPQVDVILVMSGGKI 738  
Qy 847 SEMGYSOELLARDGAFBLRTYASTEOQDAEENGVTGSGPGKEAKOMENGLVTDTSAG 906

739 SEMGYSQELLARDGAFAEFLRTVASTEQBDAAENGVTGVSFGKQAKQMGMLVTDGA 798  
 907 GKOLQOLSSSSSYSDISRHNSTAELOKAKAKKEETWKLMEADKAQTGOVKLSYWDY 966  
 799 GKOLQOLSSSSSYSDISRHNSTAELOKAKAKKEETWKLMEADKAQTGOVKLSYWDY 858  
 967 MKAIGLIFISLFIPLFMCNHNVSALASNYWLSLWTDPIVNGTQEHKVKRLSVYGALGISQ 1026  
 859 MKAIGLIFISLFIPLFMCNHNVSALASNYWLSLWTDPIVNGTQEHKVKRLSVYGALGISQ 918  
 1027 GIAVFGYSMAVSTGGILASRCLHVDLLHSILSPMSFFERTPSGNLVNRFSEKELDTVDSM 1086  
 919 GIAVFGYSMAVSTGGILASRCLHVDLLHSILSPMSFFERTPSGNLVNRFSEKELDTVDSM 978  
 1087 IPEVIMKMGSLFNIVIGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSQLKRLSEV 1146  
 979 IPEVIMKMGSLFNIVIGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSQLKRLSEV 1038  
 1147 SRSPVYSHNETLLGVSVIRAFEEQERFIHQSDLKVDENOKAYPSIVANRWLAVRLECV 1206  
 1039 SRSPVYSHNETLLGVSVIRAFEEQERFIHQSDLKVDENOKAYPSIVANRWLAVRLECV 1098  
 1207 GNCIVLFAALFAVIRSHLSLGLSVSYQVTTYNLWLVNRMSEMETNIVAVERLKE 1266  
 1099 GNCIVLFAALFAVIRSHLSLGLSVSYQVTTYNLWLVNRMSEMETNIVAVERLKE 1158  
 1267 YSETEKEAPWQIOETAPSSWPQGRVFRNRYCLAYREDLDVPLRHINVTINGGKVGIV 1326  
 1159 YSETEKEAPWQIOETAPSSWPQGRVFRNRYCLAYREDLDVPLRHINVTINGGKVGIV 1218  
 1327 GRTGAGKSLTLGLFRINESAGEEIIIDGINIAKIGLHDLRPKITIIPODPVLFSGSLRM 1386  
 1219 GRTGAGKSLTLGLFRINESAGEEIIIDGINIAKIGLHDLRPKITIIPODPVLFSGSLRM 1278  
 1387 NLDPSQYSDEBWTSLSLAHUKDFVSALPDKLDHECAEGGNSLVGQRLVCLARALLR 1446  
 1279 NLDPSQYSDEBWTSLSLAHUKDFVSALPDKLDHECAEGGNSLVGQRLVCLARALLR 1338  
 1447 KTKILVLDEATAVLETDLLIQTIRTOFEDCTVLTIAHRLNTIMDYTR 1496  
 1339 KTKILVLDEATAVLETDLLIQTIRTOFEDCTVLTIAHRLNTIMDYTR 1388

RESULT 6  
 US-09-939-853A-88  
 ; Sequence 88, Application US/09939853A  
 ; Publication No. US20040039163A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Burgess et al.  
 ; TITLE OF INVENTION: No. US20040039163A1el Proteins and Nucleic Acids Encoding Same  
 ; FILE REFERENCE: 21402-099  
 ; CURRENT APPLICATION NUMBER: US/09/939,853A  
 ; PENDING FILING DATE: 2001-08-27  
 ; PRIOR APPLICATION NUMBER: 60/228,191  
 ; PRIOR FILING DATE: 2000-08-25  
 ; PRIOR APPLICATION NUMBER: 60/267,300  
 ; PRIOR FILING DATE: 2001-02-08  
 ; PRIOR APPLICATION NUMBER: 60/269,961  
 ; PRIOR FILING DATE: 2001-02-20  
 ; PRIOR APPLICATION NUMBER: 60/277,337  
 ; PRIOR FILING DATE: 2001-03-20  
 ; NUMBER OF SEQ ID NOS: 159  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 88  
 ; LENGTH: 1528  
 ; TYPE: PRT  
 ; ORGANISM: Mus musculus  
 ; US-09-939-853A-88

Query Match 69.9%; Score 7002.5; DB 11; Length 1528;  
 Best Local Similarity 88.0%; Pred. No. 0;  
 Matches 1349; Conservative 102; Mismatches 75; Indels 7; Gaps 5;

1 MALRFGCSADGSDPLMDNNTWNTSNPDFTKCFQNTVLVWPCFYLWACFPFYLYLSRH 60  
 1 MALRFGCSADGSDPLMDNNTWNTSNPDFTKCFQNTVLVWPCFYLWACFPFYLYLSRH 60  
 61 DRGYTQMTPLNKTALGFLMIICWADLFYSFWERSRGIFLAPVFLPSPTLLGLTTLIA 120  
 61 DRGYTQMTPLNKTALGFLMIICWADLFYSFWERSRGIFLAPVFLPSPTLLGLTTLIA 120  
 121 TFLIQLERRKGVSQSGIMLTFWLVALCALAILRSKIMTALKEDAQVDLFRDITFYVYFS 180  
 121 TFLIQLERRKGVSQSGIMLTFWLVALCALAILRSKIMTALKEDAQVDLFRDITFYVYFS 180  
 181 LLLIQLVLSGSDRSPLESEITHDNPCESSASFLSRITFWMIITGLVGRVQPLESSD 240  
 181 LLLIQLVLSGSDRSPLESEITHDNPCESSASFLSRITFWMIITGLVGRVQPLESSD 240  
 241 LWSLNKEDTSEVVPVLVNNKKECAKTRKQPKVYVS- SKDPAQPKESKSKVDANEEVEA 299  
 241 LWSLNKEDTSEVVPVLVNNKKECAKTRKQPKVYVS- SKDPAQPKESKSKVDANEEVEA 299  
 300 LIVKSPQKEMNPSLFLKLYKTGPGPYFLMSFFFKAIHDLMMFSGPOILKLLIKLVNDTKAP 359  
 301 LIVKSPHDXRPSLFLKLYKTGPGPYFLMSFLYKALHDLMMFAGPKILELIINFVNDREAP 360  
 360 DWQGVYFTVLLFVTAQLOTLVLHGYFHTCFVSGWRIKTAIVGAVYRKALVITNSAKSST 419  
 361 DWQGVYFTVLLFVTAQLOTLVLHGYFHTCFVSGWRIKTAIVGAVYRKALVITNSAKSST 420  
 420 VGEIVNLSVDAQRFMDIATYINMIWSAPLOVILALYLLNLNLGSPVLAGVAVMLVMPV 479  
 421 VGEIVNLSVDAQRFMDIATYINMIWSAPLOVILALYLLNLNLGSPVLAGVAVMLVMPV 480  
 480 NAVMAKTKTQVVAHMKSKONRIKMLNEILNGIKVLKIYAWELAFKDKVLAIRQSELKVL 539  
 481 NAVMAKTKTQVVAHMKSKONRIKMLNEILNGIKVLKIYAWELAFKDKVMSIQBELKVL 540  
 540 KKSAYLSAVGHTVWCTPFLVALCTFAVYVTIDENILDAQTAFVSLALFNILRPLNIL 599  
 541 KKSAYLSAVGHTVWCTPFLVALCTFAVYVTIDENILDAQTAFVSLALFNILRPLNIL 600  
 600 PMWISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGCGGTNSITVRNATFTTWARSDPP 659  
 601 PMWISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGCGGTNSITVRNATFTTWARSDPP 659  
 660 TINGITFISIPGALVAVVGVCGCKSSLLSALLAEMDKVEGHVAIKGSVAVYVQQAQW 719  
 660 TINGITFISIPGALVAVVGVCGCKSSLLSALLAEMDKVEGHVAIKGSVAVYVQQAQW 719  
 720 DSLRENILFGQLEBPYRSVIOACALLPDLLEILPSGDRTEIGKGVNLSGGQKORVSLA 779  
 720 DSLRENILFGQLEBPYRSVIOACALLPDLLEILPSGDRTEIGKGVNLSGGQKORVSLA 779  
 780 RAVYSNADIYLFDDPLSAVDHVGKHIFENVIGPKMLKNKTRILVTHSMYSYLPQVDVII 839  
 780 RAVYSNADIYLFDDPLSAVDHVGKHIFENVIGPKMLKNKTRILVTHSMYSYLPQVDVII 839  
 840 VMSGGKISEMGYSQELLARDGAFAEFLRTVASTEQBDAAENGVTGVSFGKQAKQMGMLV 896  
 840 VMSGGKISEMGYSQELLARDGAFAEFLRTVASTEQBDAAENGVTGVSFGKQAKQMGMLV 896  
 900 MLVTDGAKQIQOLSSSSSYSGDISRHNSSTAELOKAKAKKEETWKLMEADKAQTGOVK 959  
 897 MLVTDGAKQIQOLSSSSSYSGDISRHNSSTAELOKAKAKKEETWKLMEADKAQTGOVK 955  
 960 LSVYWDYWKALIGLFIPLFMCNHNVSALASNYWLSLWTDPIVNGTQEHKVKRLSV 1018  
 956 LSVYWDYWKALIGLFIPLFMCNHNVSALASNYWLSLWTDPIVNGTQEHKVKRLSV 1015  
 1019 YGALGISOGLAVFGYSMAVSTGGILASRCLHVDLLHSILSPMSFFERTPSGNLVNRFSEK 1078  
 1016 YGALGISOGLAVFGYSMAVSTGGILASRCLHVDLLHSILSPMSFFERTPSGNLVNRFSEK 1075  
 1079 ELDTVDSNIPKEMFMSLNFNVIGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSR 1138

```
Db 1076 ELDTVDSMPQVIMKFMGSLFSGVIGAVIIILLATPIAAVILPPLGLVYFFVQRYVASSR 1135
QY 1139 QKRLSVSRSPYSHFNFTLLGVSVIRAFEEQERIHOSDLKVDENQKAYTPSVIANRW 1198
Db 1136 QKRLSVSRSPYSHFNFTLLGVSVIRAFEEQERIHOSDLKVDENQKAYTPSVIANRW 1195
QY 1199 LAVRLCVGNCIVLFAALFAVISHRSLSAGLVGLSVSYSLQVTTYNLWLVMSSEMETNI 1258
Db 1196 LAVRLCVGNCIVLFAALFAVISHRSLSAGLVGLSVSYSLQVTTYNLWLVMSSEMETNI 1255
QY 1259 VAVERLKEYSETEKEAPWIOETAPSSWPQGRVFRNYCLAYREDLDFVLRHINVTIN 1318
Db 1256 VAVERLKEYSETEKEAPWIOETAPSSWPQGRVFRNYCLAYREDLDFVLRHINVTIN 1315
QY 1319 GEKGVIQVGTGAGKSSLTGLPRINESAGEIIDGINIAKIGLHDLRFKTIIPQDPV 1378
Db 1316 GEKGVIQVGTGAGKSSLTGLPRINESAGEIIDGINIAKIGLHDLRFKTIIPQDPV 1375
QY 1379 LFGSGLRMNLDPFSQYSDDEEVMTSLELAHLKDFVSALPKLDHECAEGGENLSVGQRLV 1438
Db 1376 LFGSGLRMNLDPFSQYSDDEEVMMALELAHLKGFVSALPKLDHECAEGGENLSVGQRLV 1435
QY 1439 CLARALLRRTKILVLDDEATAAVDLETDLLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVI 1498
Db 1436 CLARALLRRTKILVLDDEATAAVDLETDLLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVI 1495
QY 1499 VLDKGIEQVGAAPSDLLQOGLFYSMAKDAGLV 1531
Db 1496 VLDKGIEVREGGAPSELLQOGRGIFYSMAKDAGLV 1528
```

## RESULT 7

```
US-09-939-853A-83
; Sequence 83, Application US/09939853A
; Publication No. US20040039163A1
; GENERAL INFORMATION:
; APPLICANT: Burgess et al.
; TITLE OF INVENTION: No. US20040039163A1e1 Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939,853A
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/228,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 83
; LENGTH: 1527
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-939-853A-83
```

Query Match 44.8%; Score 4491.5; DB 11; Length 1527;

Best Local Similarity 56.7%; Pred. No. 0;

Matches 873; Conservative 260; Mismatches 371; Indels 35; Gaps 9;

```
QY 8 SADGSDPLDMNVNTWNTSPDFTKCFONTVLMVPCFYLWACFPFFYFLVLSRDRGYIQM 67
Db 7 SEELGSKFWDNLVSVHTENPDLTTPCFQNSLLAWPCYILUWALPCVLLYLRLHCHRGYILL 66
QY 68 TPLNKTKTALGFLMTVWCADLFYFWSRSGIFLAPVFLVPTLLGITTLLATLIQUE 127
Db 67 SHLSKLMVLGLVLCWCSNADLFYSFHLVHGRAPAPVFFVPLVVGVTMLLATLIQVE 126
QY 128 RKGVOSSGIMLTFWLVALVICALIURSKIMTALKEDAQVDFRDTFTYVYFSLLIQV 187
Db 127 RLQGVQSSGLIIFWFLCVVCAIVPPRSKILLAKAEGISDPFRFTFTYIHFALVLSALI 186
```

```
QY 188 LSCFSDRSPFLSFSETIHDNPNCPRESSASFLSRITTFWMITGLIVRGYRQPLEGSDLSLNKE 247
Db 187 LACFREKPPFFSAKXNDPNPYETPSAGFLSRLFFWFTKMAIYGYRHPLDEKDLMSLKEE 246
QY 248 DTSEQVVPVLVKNWKECAKTRKQPKVVVSSKDPAPKPKSSKVDANEVEALIVKSPK 307
Db 247 DRSQWVQOOLLEAMRKQSKQTAH-----KASAAAPGK-----NAGSEDEVLLGARPP 294
QY 308 EWNPSLFVLYKTFGPYFLMSPFFKAIHDLMMFSGPOLIKLLIKFVNDTKAPDWQCVFYT 367
Db 295 R-KPSFLKALLATFGSSFLISACFKLIQDLGFINPQLLSILIRFISNPMASWNGFLVA 353
QY 368 VLLFTYACLTQTLVLHQYFHCIFVSGMRITKTAIVIGAVYRKALVITNSKSSSTVGEIVNLM 427
Db 354 GLMFLCSMMQSLILQHYHYIFVTGKFRGTGIMGYRKALVITNSKSSSTVGEIVNLM 413
QY 428 SVDAQRFDMLATYINNIWSAPLOVILALYLLMLNGLGPSVLGAVVMVLMVPMVNAVAMKT 487
Db 414 SVDAQRFDMLAPFLNLLMSAPLIQIILAIYFLWQNLGPSVLGAVFMVLLIPLNGAVAVRW 473
QY 488 KTYQVAHMKSKDNRIKLANEILNGIKVLKYAWELAFKDKVLAIRQELKVLKKSAYLSA 547
Db 474 RAFQVQKMLKDSRIKLMSEILNGIKVLKYAWELAFKDKVLAIRQELKVLKKSAYLSA 533
QY 548 VGTFTWCTPFLVALCTFAVYVITIDENNILDAQTAFVSLALFNILRFLNILPMWISSIV 607
Db 534 TTTFTWCTPFLVALCTFAVYVITIDENNILDAQTAFVSLALFNILRFLNILPMWISSIV 593
QY 608 QASVSLKRLRIFLSHEELEPOSIERRPVKDGGGTNSITVRNATFTTWARSDDPTFLNGITS 667
Db 594 QASVSLKRIQQQLSQEELDPQSVERKTISP-----YAITHTSGTFTWAOPLPTLSLDIQ 650
QY 668 IPEGALVAVGVGVGCGKSSLSALLAEMDKVEGHVAIGSVAYVPOQAIQNDLSRENLT 727
Db 651 VPKGALVAVGVGVGCGKSSLSALLAEMDKVEGHVAIGSVAYVPOQAIQNDLSRENLT 710
QY 728 FCQCLEEPYYSVIOACALLPLEILPSCDRTEIGEKGYNLSGGQKQVSLARAVYSNAD 787
Db 711 FGKALNPKRYQQTLEACALLADLEMLPGGDQTEIGEKGYNLSGGQKQVSLARAVYSNAD 770
QY 788 IYLFDDPLSADVAHVKGKHFENVIGPKMLKNKTRILVTHSMYSYLPQVDVILVMGSGKIS 847
Db 771 IYLFDDPLSADVAHVKGKHFENVIGPKMLKNKTRILVTHSMYSYLPQVDVILVMGSGKIS 830
QY 848 EMGYSQELLARDGAFAEFLRTVASTEQDABENGVTGVSQPKAKOMENGM----- 901
Db 831 EMGYPALLQRNGSPANFLCNYPADDDQHLSDSWTALEGAEDKEALLIEDTLSNHTDLT 890
QY 902 -----VTSAGKQLQRLSSSSSYSGD-----ISRHHNSTAE-LQKAEAKKETWKLMEAD 951
Db 891 DNDPVTYVYVQKQFQNLQSLASS-DGEGGQRPVPRRHGLGPSEKQVTEAKADGA--LTQEE 947
QY 952 KAQTQVQKLSVYVDYMKAIGLFISFLSIFLFCNHNVSALASNYLSLWTDPIVNGTQSH 1011
Db 948 KAAIGTVLSVFWDYAKAVGLCTTILAILLYYQSAAGAAIGANVLSAWTNDAMADSRQNN 1007
QY 1012 TKVRLSVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV 1071
Db 1008 TSLRIGVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV 1067
QY 1072 LVNREFSKELDTVDSMIPVIMKFMGSLFNVIGACIVILLATPIAAIIPPLGLIYFFVQR 1131
Db 1068 ILNCFSEKDIYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV 1127
QY 1132 FYVASSRQLKRLSVSRSPYSHFNFTLLGVSVIRAFEEQERIHOSDLKVDENQKAYTP 1191
Db 1128 FYAATSRQLKRLSVSRSPYSHFNFTLLGVSVIRAFEEQERIHOSDLKVDENQKAYTP 1187
QY 1192 STVANRWLVRLECVGNCIVLFAALFAVISHRSLSAGLVGLSVSYSLQVTTYNLWLVMS 1251
Db 1188 YIISNRWLSIGVEFGVNCVFLFAALFAVISHRSLSAGLVGLSVSYSLQVTTYNLWLVMS 1247
```



[illegible]

RESULT 9

295	R	KPSFLKALLATFGSFLISACFKUIQDLSLFINPOLLSILRFTSNWAPSWMGFLVA	353
368	VLLFTACILQTLVLHQQYFHCIFVSGMNRKIATAVGAVYRKALVITNSARKSSTVGEIVNLM	427	
354	GLMFLCSMQSLILQHYVHYFVTGQKFTGIMGVIYRKALVITNSVKRASTVGEIVNLM	413	
428	SVDARQFMDLATYINMINSAPLOVILALYLLNMLGPSVLAGVAVMVLVYVPAVNAVMAMKT	487	
414	SVDARQFMDLAPLNLMLNSAPLOIILAIYFLMQNLGPSVLAGVAVMVLIIPLMGNAVAKM	473	
488	KTYQVAHMSKONRIKLMNEILNGIKVLKLYAMELAFKDKVLAIROELKVLKKSAYLSA	547	
474	RAFQVQMKMLKDSRIKLMSEILNGIKVLKYAWEPSFLQVGEIROGELQLLRTAYLHT	533	
548	VGTFTVTCVTPFLVALCTFAVYVYIDENNILDQAOTAFVSLALFNILRPLNIPMWISSIV	607	
534	TTTFTWMCSPFLVTLITLVMVYVDPNNVLDAEKAFVSVSLFNILRPLNMLPQLISNLT	593	
608	QASVSLKRLRIFLSHEELPDSITERPVKGGGTNSITVRNAFTWARSPPTLNGITFS	667	
594	QASVSLKRIQQFLSQEELDPQSVERTISPG---YAITHSGTFTWAQDLPPTLHSLDIQ	650	
668	IPEGALVAVVGQVCGCKSLLSALLAEMDKVEGHVAIKGSVAVYVQOAWITQNDLSRENIL	727	
651	VPGKALVAVVPGVCGKSLVSALLGEMEKLEGKVMHMGSVAVYVQOAWITQNLQENVL	710	
728	FCQQLBEPYRSVIOACALLPOLEILPSGDRPTEIGKGNVLSGGQKQKRVSLARAVSNAD	787	
711	FOKALNPKRYQOTLEACALLADEMLPGGDQTEIGKGNLSGGQKQKRVSLARAVSDAD	770	
788	IYLFDDPLSADVAHVKKHIFENVIGPKMLKXKTRILVTHMSYSLPOVDVIIVMSGGKIS	847	
771	IFLLDDPLSADVSHVAKHFDHVIGPEGVLAGKTRVLVTHGISFLPQTDPIIILADGQVS	830	
848	ENGSTQELLARGOGAFBFURTYASTEQDAEENGVTGVSQPKQEKAKQMGHML-----	901	
831	ENGYPALLQRNGSFANFLCNYPADPDQGHLEDSFTALEGAEDKEALLIEDTLNSHNHDTL	890	
902	-----VTDSAGKLOQLQSSSSSVSGD-----TSRHNSTAE-LQKAEAKKEETWKLMEAD	951	
891	DNDPVTYVQKQFMQLSALSS-DGEGQGRPVRPHLGPSEKQVQVTEAKADGA-LTQEE	947	
952	KAQTOGVKLSVYDYMKAIGLFIISLFIPLFMCNHYVSALASNYNLSLWTDDPIVNGTQEH	1011	
948	KAAIGTVELSVFDYAKAVGLCTTILAIICLLYVQGSAAAIICANVWLSAWNTNDAMADSRQN	1007	

[illegible]



Db 1365 PQDPIFGSGLTNRNLDPPFCRGYSDDEDIWRTELHLSAFVSSQPTGLDFQCSGSDNLSVG 1424  
Qy 1434 QRLVCLARALKTKILVLDATAVDLETDLLQSTIRTOFEDCTVLTIAHRLNTIMD 1493  
Db 1425 QRLVCLARALKRKRVLVLDATAIDLETDLLQSTIRTOFEDCTVLTIAHRLNTIMD 1484  
Qy 1494 YTRVIVLDKGTQOEGAPSDLLQORGLFYSMADAGL 1530  
Db 1485 YNRVLVLDKGVVAEFDSPVNLIAAGGIFYGMAKDAL 1521

RESULT 11  
US-10-667-891-1  
; Sequence 1, Application US/10667891  
; Publication No. US20040171024A1  
; GENERAL INFORMATION:  
; APPLICANT: ROTH, CHARLES W.  
; APPLICANT: BREY, PAUL T.  
; APPLICANT: HOLM, INGE  
; APPLICANT: GRAILLES, MARINE  
; APPLICANT: RZHETSKY, ANDREY  
; TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS IN DROSOPHILA AND  
; TITLE OF INVENTION: ANOPHELES  
; FILE REFERENCE: 03495.0294-00000  
; CURRENT APPLICATION NUMBER: US/10/667,891  
; PRIOR FILING DATE: 2003-09-23  
; PRIOR APPLICATION NUMBER: 60/413,469  
; NUMBER OF SEQ ID NOS: 76  
; SOFTWARE: Patentin Ver. 3.2  
; SEQ ID NO 1  
; LENGTH: 1548  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-10-667-891-1

Query Match 38.5%; Score 3860.5; DB 16; Length 1548;  
Best Local Similarity 50.0%; Pred. No. 6.2e-318;  
Matches 784; Conservative 274; Mismatches 438; Indels 71; Gaps 18;

Qy 6 FCSADGSDPLDWNVTWNTSNDFTKCFQNTVLVWVPCFLWA--CFPPFYLYLRHSDRG 63  
Db 12 FCGS----TFWATETWYTNDFPTCFEQTALVMTPCAIFYWAFVDFYLYLKASL-DRN 66  
Qy 64 YQMTPLNKTALGFLMWVWADLFYFWSRSGIFLAPVF---LVSPTELLGITTLLA 120  
Db 67 -IPWKLNVSKALNGLVITALLIMALKVG-DSLEPLVLDVWGPPIKATFLL 124  
Qy 121 TFLIQLERRKGVSSGIMTLFWLVALVCAIALRSKI-----MTALKEDAQVD 168  
Db 125 FIFILNRKYGVQTTGCGQIFWFLTLVLSIPRCRTEVRLDAERQKILNSQQPSEQDFSW 184  
Qy 169 LPRDITFYVYFSLLIQLVLCFSDBSLPSETIHDNPNCPRESSASFILSRITFWITGLI 228  
Db 185 EYQFVSFFIFFFTSMLTLCFADGMPRTQYQRGENSEIPLSASFILSRITYQWFDKMA 244  
Qy 229 VRGYRQPLGSDLSLNLKEDTGEQVVPVLVKNWKECAKTRKQPVKVVYSSKDPAPKRES 288  
Db 245 LKGYRNPLEKLDLRLPDQSCSEVNPPIFAHWNQVRKNYQKARV-----EPKAQFNG 300  
Qy 289 SKVDANEVEALIVKSPQKMNPSLFKVLKTYFGPYFLMSFFPKAIDHLMFSGPQILKL 348  
Db 301 NVTFENPHGE---KNGRKKGMASIMPIYKSGGVFLFGALMKLFTDLTFAQVQLSL 356  
Qy 349 LKFNVDTKA-PDWQGYFTVLLFVTACLOTLVHQYFHCFTVSGWRKRTAVIGAVYRKA 407  
Db 357 IISFVEAQDAEPKGYLAVLLFVLAQAQTFILGQYFHRMFIVGIRINTALINAIYRKA 416  
Qy 408 LVITNSARKSSTVGETVNLMSVDAORFMDLATYINMWSAPLQVITALLVLLNLNLPSPVL 467  
Db 417 LRISNTKTESVGETVNLMSVDAORFMDLATYINMWSAPLQVITALLVLLNLNLPSPVL 476  
Qy 468 AGVAVMVLMPVNVAMKTKTYQVAHMKSKONRIKLANEILNGIKVLKLYAWELAFKDK 527

Db 477 AGLAVMILIPVNGVIASRIKTYQIRQMKYKDERVKLMEVLSGIKVLKYAWERSFEKQ 536  
Qy 528 VLAIHQEELKVLKSAVLSAVCTFTWCTPPLVALCTFAVYVTDENNILDAQTAQVSLA 587  
Db 537 VLDIHKDKIATURSTAYLNAGTSFLWSCAPPLVSVLTATYVLIIDENNVLDAKTATFVSU 596  
Qy 588 LFNILRFLNLPVMISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGGTNSITVR 647  
Db 597 LFNILRFLNLPMLITNLVQTVSVNRINKFNLSELDPSV---LHDSKPKHPMSIE 652  
Qy 648 NATFTWASDPPTLNGITFTSPEGALVAVGVGVGCKSSLLSALLAEMDKVGHVAIKGS 707  
Db 653 NGEFSW--GDEITLRNINI EVKKGSLVALVGTVGSGKSVVQAFLGEMEKLAGVNVTVGK 710  
Qy 708 VAYVPOQAWIQNDSURENIFGQLEEPYRSVIOACALLPDLLEILPSSDRTEIGKGVN 767  
Db 711 LAYVPOQAWIQNATVRDNLFGQTYDRKRYNKVIDACALRADIDILSAGDLTEIGKGIN 770  
Qy 768 LSGGQKQVRSARAVYSNADIYLFDDPLSAVDHVGKHFENVIGPKGMLKNKTRILVTH 827  
Db 771 LSGGQKQVRSARAVYSNADIYLFDDPLSAVDHVGKHFENVIGPKGMLKNKTRILVTH 830  
Qy 828 SMSYLPQVDVIVMSGGKISEMGSQELLARDGAPAEFLRTYASTEQEODAEEN-----881  
Db 831 GVTFLPQVDSIVYIKVGEISESGTDPQLVKNKGAFADFIQHLQEGNEEBEELNQIKRQI 890  
Qy 882 ---GVTGVSQPGKEAKQWENGMLVDS-----AGQLQORLSSSSSYSGDISR 926  
Db 891 SSTADVPCELLGTVEKAIKLARTESLSDSISVTSADSLMGGGGSLRRRTKQDS-----943  
Qy 927 HNSTAEELQKAEKEETWKLMEADKAQTQGVKLSVYMDYMKAIGLFISFLIFMCMNH 986  
Db 944 -HDSVASAASLKKQKQEVSGKLIETESQTGGVEFAVYKHYIKSGIFLSVATLVNLFVQ 1002  
Qy 987 VSALASNWLSLWTDPIVNGTQEHKTVRLSVYGALGISQGIATVFGYSNAVIGGILASR 1046  
Db 1003 AFQIGSNLWLTQWANDQNVANDTGLRDMYLVGYGAFGQVLSKYLGLALAIIGGLHCSM 1062  
Qy 1047 CLHVDLLSHLSRSPMSFFERTPSGNLVNRFSKELDTVDSMIPEVIMKPMGSLFNVIACI 1106  
Db 1063 NVFNKLLNTGLKPMELFDTPLGLRILSKYKQDVTVDVSLPAITVQLLNTCFGLVATTIV 1122  
Qy 1107 VILLATPAAIIIPPLGLIYFFVQFYVASSRQLARLESVSRSPYSHFNELLLGVSVTR 1166  
Db 1123 VLSLTPFLAVIVPIAFIYFAQFYVATSRQLARLESVSRSPYSHFSETVTGASTIR 1182  
Qy 1167 APEQERTIHOSDLKVDENQKAYPSIVANRWLAVLECVGNCIVLPAALPAVIRSHLS 1226  
Db 1183 AYNVGDRIEESDAKVDKNQVCKYPSVIANRWLAIKLEWVGNLIILFASLFAVLGGQT-N 1241  
Qy 1227 AGLVGLSVYSLOVTTYNLWLRMSMETNTVAVERLKEYSETKEAPWQI-QETAPPS 1285  
Db 1242 PGLVGLSVYALQVQTNLWLRMSDDETNIIVSVERIKEYGETKQEAPEWELQKQKPK 1301  
Qy 1286 SWPQVRFEFRNYCLIRYREDLDFVLRHINVTINGCKEYGVIGRTGAGKSSLLTGLFRINE 1345  
Db 1302 NWPQGRVEFQNFQVRYREGDLVLRGVSNFOGGEKVGIVGRTGAGKSSLLTGLFRINE 1361  
Qy 1346 SAEGEIIIDGINIAKIGLHDLRFKTIIPQDPVLPFSGSLRNMNLDPPFSQVSDSEVTSLEL 1405  
Db 1362 AAGGRISIDGVDIASMGLMLRSRLTIIPQDPVLPFSGSLRINLDPFEIKTDDDEIKWALEL 1421  
Qy 1406 AHLKDFVSALPDKDHECAEGENLSVGORQLVCLARALLRKTLLVLDDEATAAVDLETD 1465  
Db 1422 SHLSFVLSLAAGLHNEIABEGENLSVGORQLVCLARALLRKTLLVLDDEATAAVDLETD 1481  
Qy 1466 DLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVLDKGEIOEYAGAPSDLLQO-RGLFYSM 1524  
Db 1482 DLIOKTIETPEKCTVLTIAHRLNTILSDKVIIVLDKQIIEFASFTELLDNPKSAFYSM 1541  
Qy 1525 AKDAGLV 1531  
|||||

```

Db      1542 AKDANLV 1548

RESULT 12
US-10-667-891-4
; Sequence 4, Application US/10667891
; Publication No. US20040171024A1
; GENERAL INFORMATION:
; APPLICANT: ROTH, CHARLES W.
; APPLICANT: BREY, PAUL T.
; APPLICANT: HOLM, INGE
; APPLICANT: GRAILLES, MARINE
; APPLICANT: RZHETSKY, ANDREY
; TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS IN DROSOPHILA AND
; TITLE OF INVENTION: ANOPHELES
; FILE REFERENCE: 03495.0294-00000
; CURRENT APPLICATION NUMBER: US/10/667,891
; CURRENT FILING DATE: 2003-09-23
; PRIOR APPLICATION NUMBER: 60/413,469
; PRIOR FILING DATE: 2002-09-26
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: Patentin Ver. 3.2
; SEQ ID NO 4
; LENGTH: 1505
; TYPE: PRT
; ORGANISM: Anopheles gambiae
US-10-667-891-4

Query Match      36.0%; Score 3603.5; DB 16; Length 1505;
Best Local Similarity 48.0%; Pred. No. 5.9e-296;
Matches 750; Conservative 280; Mismatches 444; Indels 87; Gaps 22;

QY      1 MALRGCASADGSDPLDWNVTWNTGNPDFTKCFQNTLVWVPCFVLMACFPFYLY-LSR 59
Db      1 MTFDEFCG-----PFDGFEFVDVNDPNLTFQFQVILQWPCLEFV-FSIYDFKITE 55
QY      60 HDRQVIQMTPLNKTATGFLFLMWVCADL-FYSFWSRSGIFLAPVFLVSPTLGIGITL 118
Db      56 SKYRIDPWNWNLRSMLVFLMLCMWIDLGMVGYQDBEQLYDQILTAFAVNALAYIDL 115
QY      119 LATELIQLERRKGVQSSGIMLTFWLVVALVCAIILRSKIM-----TALKEDAQVDL--F 170
Db      116 LV--LJFWRKYGVRISGTMFMFWRPFGLIQLRTEVMENDKRPNAIGSGDVTDFWEY 173
QY      171 RDIETVYVFSLLIQLVSCFSDRSPLFSETHDNPCESSASPLSRITFTFWITGLIVR 230
Db      174 QVSVYILQSLICLMLVLELPDKEPSFYYPKAAKPNELRSFFSKLLFLHFDPAFWK 233
QY      231 GYRQPLEGSDLSLNKEDTSEQVPVLVKNKKCEKATRKQPKVYVYSSKDPAPKSSK 290
Db      234 GFRNPLTNDMYDINPQDSARELVPPFDKYWKISVEKGRKQKQ-----ASDRKAGK 284
QY      291 VDANEVEALIVKSPQKEMNPSLFKVLKTF-----GPYFLMSFFPKAIHDLMMFSGQIL 346
Db      295 PDID-----YKPHSFNSGSLVYTMTRAVGGPFWFAGMLQLAISGL-QFASPLYM 332
QY      347 KLLIKFVNDTKAPDQGGFYTVLLFVTACQLTLVLHQQVFHICFVSGMRITKATVAVYRK 406
Db      333 QELMAVI-AFDGPLWKGFLLTFLGASLLGLFNGQVLYFTFLSGFRIRGLISALYRK 391
QY      407 ALVITNSARKSTVGEIVNLMSVDAQRPMDLATYINNIWSAPLOVIALYALYLLWNLGFSV 466
Db      392 ALRISSAAKDTTVGEIVNLMAVDAQRPFELTSLYHLMSALLIIGLCVFLLYDILGPV 451
QY      467 LAGVAVMVLVMPVNAVMAKTKTYQVAHMKSKDNRIKLMNBLNGIKVLKLYAWELAFKD 526
Db      452 FAGLGVMILMTFVSGVAAKLKTHQVQMKLKDERVKKMKNEILGIIKVLKLYAWEPFQD 511
QY      527 KVLAIROBELKVLKKSAYLSVNGTFTWCTPFLVALCTFAVYVITDENNILDQAQFVSL 586
Db      512 SILNVRDEEVGILKMAVYAGIETFTIAPFLVTVLSFVAVYVILIDENNVLDPQTAQFVSL 571
QY      587 ALFNILRFPPLNLPWIVISSIVQASVSLKRLRIFILSHHEELPDSIERPVPKGGGTNSITV 646

```

RESULT 13  
 US-10-667-891-3  
 ; Sequence 3, Application US/10667891  
 ; Publication No. US20040171024A1  
 ; GENERAL INFORMATION:

```

; APPLICANT: ROTH, CHARLES W.
; APPLICANT: BREY, PAUL T.
; APPLICANT: HOLM, INGE
; APPLICANT: GRAILLES, MARINE
; APPLICANT: RZHETSKY, ANDREY
; TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS IN DROSOPHILA AND
; TITLE OF INVENTION: ANOPHELES
; FILE REFERENCE: 03495-0294-00000
; CURRENT APPLICATION NUMBER: US/10/667,891
; CURRENT FILING DATE: 2003-09-23
; PRIOR APPLICATION NUMBER: 60/413,469
; PRIOR FILING DATE: 2002-09-26
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: Patent in Ver. 3.2
; SEQ ID NO 3
; LENGTH: 1499
; TYPE: PRT
; ORGANISM: Anopheles gambiae
; US-10-667-891-3

Query Match          36.0%; Score 3601.5; DB 16; Length 1499;
Best Local Similarity 47.9%; Pred. No. 8.6e-296;
Matches 744; Conservative 283; Mismatches 448; Indels 79; Gaps 20;

QY 1 MALRGFCSADGSDPLMDWNTWNTNSPDTKCFQNTVLVWPCFYLWACFPFFYFLYLSRH 60
DB 1.MTFEEFCGG-----PFWDDDLTWRDEDDPLTCFQRIQWLTQCFPLFVFSMYEVLRIVTS 56

QY 61 DRGYQMTPLNKTALGFLWVWADLIFYSWERSRGIFLAPVLPVSPITLLGIT---- 116
DB 57 RYRDLPMWNTNMTFTALMVMGWVDL-----GVGLANDWIES--LLSMLPSLP 105

QY 117 -TLATFLQLERRKGVSSGIMLTFWLVALCALAILRSKMT-ALKEDAQVDL--FRD 172
DB 106 FOIMAMALYFFYKYGIRSTGTMFIWFLKAFGIIQMTTEAMLHDVRSGTGDPFAEQF 165

QY 173 ITFYVFSLLLQLVLSFSDSPLEFSETHIDNCPDESSASFSLRITFTWITGLIVRGY 232
DB 166 VSYTIQYTPVCCVLLLELPDREPRYSEWAKLNPNELRSLFFSRLFLYLPDSYAWRGF 225

QY 233 RPLGSDLWSLNKEDTSQVVPVLVQWKECAKTRKQPVKVYSSKDPQPKESSKVD 292
DB 226 RKLPTDDMDYLNPDSTRALVPPDPKYYWSEVGRKKQI-----AADKKAGKTN 276

QY 293 ANEEVEALIKVSPQEWNPSPFLKLYKTF-GPYFLMSFPFKAHDLMMFSGPQILKLTIK 351
DB 277 -----LVYKPNATNGSVLPAMVKYGGFFWAGMLQFALSGL-QFASPYLMQEIWA 327

QY 352 FVNDKADPWQGYFTVLLFVTTACIQLTVLHYFHCFSVGMRIKTAIVIGAVYRKALVIT 411
DB 328 VI-ALDGGPFWKGMIIITLGLFLTLALLFNGQYFHTFLVGPRIKTLISAIYRKALRIS 386

QY 412 NSAKSSTVEATVNLMSVDAQFMDLATVNNMWSAPLOVILALVLLMLNLPSPVLGVA 471
DB 387 SFPAKDDTTVEIVNLMAVDAQFFELTSYHLVLSAPLIACIYLLYELGPAVFAVLG 446

QY 472 VNLVMPVNAVMAKTKTYQVAHMKSKONRIKLMNEILNGIKVLKLYAWELAFKDKVLAI 531
DB 447 VNVIMPIPTGATMRDLQVQOMKIKOERVKKNNEILGKIKVLYAWEPFQDTVTIV 506

QY 532 ROEELVKLSAYLSAAGFTVWCTPFLVALCTFAVYVTTIDENNILDQATFVSLALFNI 591
DB 507 RNEELDVLKSAAYGAGTYFVWMAFPLVTLASFAYVYMIDENVDLPQATFVALALFNI 566

QY 592 LRFPLNLPMTVSSIVQASVSLKRLRIFLSHLELEPDSITERPVKDGGSITVRYNATF 651
DB 567 LRFPLAMPFMTITFAQAWWSIKRIDKFNWSELDPNVTHNK-----SENALVKGDTGF 621

QY 652 TWASDDPPPTINGITTFISPEGALVAVVGVQCGKSSILLSALLAEMDKVEHVAIKGSVAVY 711
DB 622 SWG-DDAFTLKNILNLRGKLSAVVGGVGTGKSSILSALLGEMKMGKSVNTDGSIAVY 680

QY 712 PQAWIQNDSLRENILFGCOLBEPYRSVIQACALLPDLLEILPSGDRTEIGEKGYNLGG 771

```

RESULT 14

```

US-10-363-112-4
; Sequence 4, Application US/10363112
; Publication No. US20040091964A1
; GENERAL INFORMATION:
; APPLICANT: THE AUSTRALIAN NATIONAL UNIVERSITY
; APPLICANT: BOARD, PHILLIP
; APPLICANT: HARRIS, MATTHEW
; TITLE OF INVENTION: MODIFIED PROTEINS, ISOLATED NOVEL PEPTIDES, AND USES THEREOF
; FILE REFERENCE: 007643-0302189
; CURRENT APPLICATION NUMBER: US/10/363,112
; CURRENT FILING DATE: 2003-11-03
; PRIOR APPLICATION NUMBER: PCT/AU01/01093
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: 60/229,663
; PRIOR FILING DATE: 2000-08-31
; NUMBER OF SEQ ID NOS: 62

```

```

DB 681 PQAWIQNATLRLNLFGRPFDOAKYDKVIECCALRPDLEMLPGDPTTEIGEKGINLGG 740
QY 772 OKQVSLARAVYSNADIVLFDPLSADVAHVGKHFENVIGPKMLKKNKTRILVTHSMYS 831
DB 741 QKQVALLARAVYADSEVYLFDDPLSADVAHVGKHFKEVIGSGMLVGRSLRLVTHGISF 800
QY 832 LPQVDVIVMSGGKISEMSYOELLARDGAFAEFLRTYASTQEODAEENGVTGVSGPOK 891
DB 801 LPFVEEIPVMDGVESESGSQELLDQKGAFAEFLTQHI---QEMDDEDE-----D 848
QY 892 EAKQWENGMLVTDTSAGKQLQRLSSSSSSYSGDI-----SRHNSSTAELOKAE 938
DB 849 ELKLIQEA-LKDGEAKTIVQRAMSTRSQSSNGSVRKKRVSRASRNSNKPRAVEQTV 907
QY 939 AKKEETWKLMEADKATGOVKLSVYDYMKAIGLFIISFLSIFLFCNCHVSALASNYWLSL 998
DB 908 AQOSSA-TLIEKESATGAVGVVYIKYFGIGLWGFWSFFSVINOGASIVAINWLTLD 966
QY 999 WTDDPIVNGTQHTKTVLSVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVY 1058
DB 967 WSEDEPAATDPSVRDMLGVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVY 1026
QY 1059 SPMSPEPTPSGNLVNRESKELDTVDSMIPEVIMFMSGLNVCACIVILLATIAAII 1118
DB 1027 MEMSFEDTTLGRINNRFSKVDVVDNIPLOSIRAWLLMFFNVIGVYVYVYVYVYVYVY 1086
QY 1119 IPLGLIYFFVQRFVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVY 1178
DB 1087 VPALVIVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVY 1146
QY 1179 DLKVDENQAYVPSIVANRWLAVERLCVNCIVLFAALFAVISRHSLSAGLVGSVYSL 1238
DB 1147 EQRVDYNQLTYSYPSIIANRWLAVERLCVNCIVLFAALFAVISRHSLSAGLVGSVYSL 1206
QY 1239 QVTYVNLWVMSSEMETNIIVAVELKESYSETEKAPHQIQETAPPSPPQVGRVFRNY 1298
DB 1207 QISATSLFVMTAEVETNIIVAEILEEYTVLPREAEWQLGHV--DKAMPVEGKVEFDY 1264
QY 1299 CLRYREDLDFVLRHNVINGEKVIGVGRGAGKSSLTGLGFRNESAEGIIIDGINI 1358
DB 1265 QIRVREGDLVIRGSLNVRGEGIKVIGRGTAGKSSLTGLGFRIVEAAGGIIIDGLDI 1324
QY 1359 AKIGLHDLRFKITTIPQDPVLFSGSLRNMNLPFSQYSDSEVWTSLELAHLKDFVSALPK 1418
DB 1325 SKWGLHLQRLTIIIPQDPVLFSGTLRANVPFKSYSDLVWKALESLHLKTFVKGAAAG 1384
QY 1419 LDHECAEGENLGVQORQLVCLARALLRKTLLVLEATAVADLETDLLIQTSTIRTPED 1478
DB 1385 LDHEIAENGSLVQORQLICLARAVALRKTLLVLEATAVADLETDLLIQTSTIRTEPAD 1444
QY 1479 CTVLTIARLNTIMDYTRVIVLDKGEIOEYGAPSDLLQOR-GLFYFSMAKDAGLV 1531
DB 1445 CTILTIARLNTILSDRVLVDKGLVAECDSQPNLLANRESIFPGMAKNAGIV 1498

```

SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 4  
 LENGTH: 1542  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-10-363-112-4

Query Match 35.9%; Score 3600; DB 15; Length 1542;  
 Best Local Similarity 47.6%; Pred. No. 1.2e-295;  
 Matches 747; Conservative 287; Mismatches 462; Indels 72; Gaps 20;

QY 3 LRGFCSADGSDPLMDNWNVTWNTS-----NPDTKCFQNTLVVWPCFYLVWACFPFFYL-- 55  
 DB 2 LKFCN-----STWNSFLDPSPEADPLCEQVTLVWVPLGLFLLAPWQLLVH 51

QY 56 YLSRHRDGYIOMTPLNKTALGFLWVWADLFYSFWERGRGIFLAPVFLVSPITLGI 115  
 DB 52 YSRTRKRSSTTKLYLAK-QVFGVLL-ILAAELALVLTEDSQOATVPVAVRYTNP-SLY-L 108

QY 116 TTLATFLQLERRKGVQSSGIMLT-FWLVALVCAILAIRSKIMTALKEDAQVDLFRDIT 174  
 DB 109 GTWLLVLLLOYSRQWCVQNSWFLSLFWILSLILCGTFQFQTLIRTLLOQD-NSNLAYSCL 167

QY 175 FVYFSLLLIQLVSCFDRSPLFSETIHDNPNCPSSASPLSRITFWITGLIYRGYRQ 234  
 DB 168 FTISYGFQILILTFASSEN-----NESSNPNSSIASFLSSITYSYWDSIILKGYKR 219

QY 235 PLEGGDLWLNKEDTSEQVVPVLVKNWKECAKTRK-----OPKVVVYSKD 281  
 DB 220 PTLTEDVWEDEMTKTLVSKFETHMKRELQARRALQRRQEKSSQNSQNGARLPGLNKN 279

QY 282 PAQPKESSVDANEEVEALIVKSPQEWNPS--LFKVLYKTFPGVPLFMSFPFKATHDLM 339  
 DB 280 QSOQDALV---EDVEKKKSGTKKQVPSKWLKALFKTFYVWLLKSLFLKLVNDIFT 336

QY 340 FSGPQLKLLIKFVNDTKADPDQGYFYVLLVFWTACLOFLVHQYHFCFVSGMRTAV 399  
 DB 337 FVSPQLKLLISFASDRDTYLMIGVLCAILLFTAAALIQSFCLQCYFCFLGKLVKVRTAI 396

QY 400 IGAVYRKALVITNSARKSTVGEIVNLMSVDAQREFMDLATYINMTWASAPLQVILALYLL 459  
 DB 397 MASVVKKALTSLNARKEYTVEGTNLMSVDHQAQKMDVTNFMHLMWSVQLVLSIFFL 456

QY 460 NLGPSVLAVGVAVMVLVAVNAVMAMTKTQVAHMKSKDNRIKLMNBLNGIKVLKYA 519  
 DB 457 RELGPSVLAVGVAVMVLVAVNAVMAMTKTQVAHMKSKDNRIKLMNBLNGIKVLKYA 516

QY 520 WELAFKDVLAITROELKVLKSAVLSAVCTFTWCTFPLVALCTFAVYVITDENNLDA 579  
 DB 517 WEPSRQVQNLKELKLLAFSOLQCVFVFLVLTPLVSVVTVFVYVLDVSDNLLDA 576

QY 580 QTAFLVSLALFNILRPLNIPWISSIVQASVSLKRLRIFLSHELEPDSITERRPVKDG 639  
 DB 577 QKAFSTITLNLRLPLMLPMWISSMLOASVTERLEKYLGGDDLDTSARH-----DCN 632

QY 640 GTNSTVTRNATFTWARSDDPTLNGITFISPEGALVAVGVQVCGKSSLLSALLAEMDKVE 699  
 DB 633 FDKAMQFSEASTWEHDESEATVRDNLDMAGQLVAVIGVPGVSGKSSLSISAMLGEMNVH 692

QY 700 GHVATKGVAVVPOQAMTQNSLRENILFGCLQLEPPYRSVIOACALLPDLLEILPSGDR 759  
 DB 693 GHITIKGTAVPOQSMITQNGTIKDNILFGFEKRYQQVLEACALLPDLLEILPGGDLA 752

QY 760 EIGKGNLGGQKQKRVSLARAVYNSADIYLFDDPLSAVDHVGKHFENVIGPKGMLKN 819  
 DB 753 EIGKGNLGGQKQKRVSLARAVYNSADIYLFDDPLSAVDHVGKHFENVIGPKGMLKN 812

QY 820 KTRILVTHSMYLPQVDVIVVMSGGKISMGSGYQELLARDGAPAEFTRTY-ASTEQEQDA 878  
 DB 813 KTRILVTHSMYLPQVDVIVVMSGGKISMGSGYQELLARDGAPAEFTRTY-ASTEQEQDA 872

QY 879 -----FENGVTGSGGKEAKQWENGMLVTDGAGLOLQOLSSSSYSODISRHNST 931  
 DB 879 -----FENGVTGSGGKEAKQWENGMLVTDGAGLOLQOLSSSSYSODISRHNST 931

DB 873 TVHDSEREDDYGILISSVEEIP--EDAASITMRRENSFRRTLSSRSSRNGRHLKSLRNS 930  
 QY 932 AELQKAEAKKEET-----WKLMEADKAQGTQVYKLSVYMDYMKAIGLFISFLIFLFCMCH 986  
 DB 931 LKTRNVNSLKEDEELVKQGLIKKEFIETGKVKFSIYLEYLQAIGLFSFFILLAFVNS 990

QY 987 VSALASNYWLSLWT--DDPIVNGTQ---EHTKVLRSVYGALGISQIAGVGYSMVSIIGI 1042  
 DB 991 VAFIGNLWLSAWTSKIFNSTDYPASQORDRMVGYGALGLAQGIFVFAHFWSAFGPFV 1050

QY 1043 LASRCLHVDLHLSILRSPMSFERTPSGNLVNRPSEKELDTVDSPMPEVIMKPMGSLFNVI 1102  
 DB 1051 HASNLHQLNLLNLRAPMRFDPTTGRIVNRFAGDLSITVDDTLQSLRSWITCFGLII 1110

QY 1103 GACIVILLATPTAAIIIPPLGLIYFFVQRFYVASSRQLRLESRSRSPVYSHFNELTGV 1162  
 DB 1111 STLVNMCATPVFTIIVIPGLIYVQVQFVSTSRQLRRLDSVTRSPYSHFSETVSG 1170

QY 1163 SVIRAFEEQERFIHOSDLKVDENQKAYPSIVANRWLAVRLECVGNCIVLFAALPAVISR 1222  
 DB 1171 PVIRAFEEHQERFLKHEERIDTNQKCVFSWITSNRWLAIRLELVGNLTVFVSALMMVYR 1230

QY 1223 HSLSAGLVGLSVYSLOVTVTLNMLVRMSSEMTNIVAVERLKEYSETEKAPWIOETA 1282  
 DB 1231 DTLSGDTVGVLSNALNITQTLNMLVRMTSEIETNIVAVERTYTKVENEAPW-VTKR 1289

QY 1283 PPSWQVQVRFNRYCLRYREDLOFLVRHNVNTINGGKVGIVGRTGAGKSSLTGLFR 1342  
 DB 1290 PPDWPSKGIQPNYQVRYRPELDLVRGITCDIGSMEKIVGVGRTGAGKSSLTCLFR 1349

QY 1343 INESAEGSIIIDGNIAGIHLDRFKITIIIPQDPVLFGSLRMLNLDPSOYSDVEEWT 1402  
 DB 1350 ILEAAGQIIDGVDIASIGLHDLREKLTIIIPQDPILFGSLRMLNLDPSOYSDVEEWT 1409

QY 1403 LELAHLKDFVSALPDKLDHECAEGENLSVQORQLVCLARALLRKTKILVLDEATAVDL 1462  
 DB 1410 LELAHLKDFVSALPDKLDHECAEGENLSVQORQLVCLARALLRKTKILVLDEATAVDL 1469

QY 1463 ETDDLIOSTIRTOEDCTVLTIAHRLNTIMDYTRVIVLDKGEIOEYGAPELLOOGLFY 1522  
 DB 1470 ETDNLIOSTIRTOEDCTVLTIAHRLNTIMDYTRVIVLDKGEIOEYGAPELLOOGLFY 1529

QY 1523 SMAKDAGL 1530  
 DB 1530 FMAKEAGI 1537

RESULT 15  
 US-10-252-155-4  
 ; Sequence 4, Application US/10252155  
 ; Publication No. US20040068096A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bristol-Myers Squibb Company  
 ; TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS IN ORGANIC ANION TRANSPORT  
 ; TITLE OF INVENTION: MULTI-DRUG RESISTANT PROTEINS  
 ; FILE REFERENCE: D0152 NP  
 ; CURRENT APPLICATION NUMBER: US/10/252,155  
 ; PRIOR FILING DATE: 2002-09-20  
 ; PRIOR APPLICATION NUMBER: US 60/324,172  
 ; PRIOR FILING DATE: 2001-09-21  
 ; PRIOR APPLICATION NUMBER: US 60/333,700  
 ; PRIOR FILING DATE: 2001-11-27  
 ; NUMBER OF SEQ ID NOS: 783  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 4  
 ; LENGTH: 1545  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-10-252-155-4

Query Match 35.9%; Score 3600; DB 15; Length 1545;  
 Best Local Similarity 47.6%; Pred. No. 1.2e-295;  
 Matches 747; Conservative 287; Mismatches 462; Indels 72; Gaps 20;



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2005, 10:57:19 ; Search time 139.718 Seconds  
(without alignments)  
5389.603 Million cell updates/sec

Title: US-10-665-283-8

Perfect score: 10016

Sequence: 1 MALRGFCGADGSDPLMDWNV.....RSVAVAKPKFSPDLSL 1947

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_15Dec04.\*  
1: Geneseq1980s.\*  
2: Geneseq1990s.\*  
3: Geneseq2000s.\*  
4: Geneseq2001s.\*  
5: Geneseq2002s.\*  
6: Geneseq2003as.\*  
7: Geneseq2003bs.\*  
8: Geneseq2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7860	78.5	1531	2	AAW57486 Human MRP
2	7860	78.5	1531	2	AAW74471 Human mul
3	7860	78.5	1531	2	AAW99894 Human mul
4	7860	78.5	1531	3	AAV55799 Human mul
5	7860	78.5	1531	3	AAV78873 Multidrug
6	7860	78.5	1531	5	ABG61810 Prostate
7	7860	78.5	1531	6	ABM35012 Cancer ba
8	7860	78.5	1531	6	ABD20865 MRP1 base
9	7860	78.5	1531	7	ADB87954 Human UGT
10	7860	78.5	1531	7	ADB96937 Human MDR
11	7860	78.5	1531	7	ADB92128 Human MDR
12	7860	78.5	1531	7	ADD44764 Human Pro
13	7860	78.5	1531	8	ADN97111 MRP1 prot
14	7856	78.4	1530	7	ADN95929 Human BEC
15	7849	78.4	1531	2	AAV54928 Multidrug
16	7849	78.4	1531	2	AAV93153 Multi-dru
17	7849	78.4	1531	2	AAW57485 Human mul
18	7849	78.4	1531	2	AAW74470 Human mul
19	7849	78.4	1531	3	AAW99893 Human mul
20	7849	78.4	1531	3	AAV55798 Human mul
21	7849	78.4	1531	3	AAV78872 Human mul
22	7849	78.4	1531	3	AAV03582 Multidrug
23	7810	76.0	1489	2	AAV96952 Multi-dru
24	7587.5	75.8	1482	8	ADN03902 Antipsori
25	7587.5	75.8	1482	8	ADP23422 PRO poly

26	7572	75.6	1481	5	ABP52108 Homo sapi
27	7560.5	75.5	1482	2	AAV96953 Multi-dru
28	7500.5	74.9	1472	2	AAV96954 Multi-dru
29	7230.5	72.2	1530	7	ADP56611 Bovine MR
30	7194	71.8	1417	2	AAV96955 Multi-dru
31	7037	70.3	1388	7	ADJ69912 Human hea
32	7002.5	69.9	1528	2	AAW57487 Murine mu
33	7002.5	69.9	1528	2	AAW74472 Mouse mul
34	7002.5	69.9	1528	2	AAV99895 Mouse mul
35	7002.5	69.9	1528	3	AAV55800 Murine mu
36	7002.5	69.9	1528	3	AAV78874 Murine mu
37	6916	69.0	1372	2	AAV98211 Multi-dru
38	6911	69.0	1345	2	AAV96956 Multi-dru
39	4491.5	44.8	1527	5	AAU91309 Human pro
40	4491.5	44.8	1527	5	ABP52110 Homo sapi
41	4491.5	44.8	1527	6	ABR58639 Human can
42	4491.5	44.8	1527	7	ADP54387 Human pro
43	4491.5	44.8	1527	7	ADN40024 Cancer/an
44	4491.5	44.8	1527	8	ADP24088 PRO poly
45	4487.5	44.8	1527	2	AAV43543 A human M

## ALIGNMENTS

RESULT 1  
AAW57486  
ID AAW57486 standard; protein; 1531 AA.  
XX  
AC AAW57486;  
XX  
DT 14-AUG-1998 (first entry)  
XX  
DE Human MRP variant ltpgpa (Lei/Pgpa).  
XX  
KW Multidrug resistance-associated protein; MRP; tumour; human; variant;  
KW multidrug resistance; MDR; leishmania P-glycoprotein; ltpgpa; Lei/Pgpa.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 685 /label= L685S  
FT /note= "wild-type Leu is replaced by Ser"  
FT Misc-difference 1282 /label= R1282A  
FT /note= "wild-type Arg is replaced by Ala"  
XX  
PN US5766880-A.  
XX  
PD 16-JUN-1998.  
XX  
PF 05-JUN-1995; 95US-00463092.  
XX  
PR 27-OCT-1992; 92US-00966923.  
PR 08-MAR-1993; 93US-00029340.  
PR 26-OCT-1993; 93US-00141893.  
PR 20-MAR-1995; 95US-00407207.  
XX  
XX (TOOH ) UNIV QUEENS KINGSTON.  
XX  
XX Cole SP, Deeley RG;  
XX WPI: 1998-361687/31.  
XX N-PSDB; AAV31498.  
XX DNA encoding protein associated with multi-drug resistance - useful for  
XX as probe for identifying multi-drug resistant tumour cells.  
XX Claim 1; Col 67-78; 82pp; English.  
XX This represents a variant of the human multidrug resistance-associated  
XX protein (MRP). This natural variant is a leishmania P-glycoprotein related

CC molecule lcpGpA (Lei/PgpA). The human and murine MRP nucleic acid  
 CC molecules can be used as probes for identifying multidrug resistant  
 CC tumour cells by hybridisation to mRNA from tumour cells. The antisense  
 CC nucleic acid can be used to reverse multidrug resistance (MDR). A  
 CC recombinant expression vector containing the MRP nucleic acid molecules  
 CC operatively linked to at least one regulatory sequence can be used to  
 CC transform a host cell to produce a recombinant MDR-associated protein  
 XX  
 SQ Sequence 1531 AA;

Query Match 78.5%; Score 7860; DB 2; Length 1531;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGCSDGSDPLMDNNTNTNSPDFTKCFQNTVLWVPCFYLWACFPFYLYLSRH 60  
 DB 1 MALRGCSDGSDPLMDNNTNTNSPDFTKCFQNTVLWVPCFYLWACFPFYLYLSRH 60

QY 61 DRGYTQMTPLNKTALGFLMWVCWADLFYFWSRSGIFLAPVPLVSPPTLGLITLLA 120  
 DB 61 DRGYTQMTPLNKTALGFLMWVCWADLFYFWSRSGIFLAPVPLVSPPTLGLITLLA 120

QY 121 TFLIQLERRKGVQSSGIMLTFWLVLCALAILRSKIMTALKEDAQVDFRDITFYVYFS 180  
 DB 121 TFLIQLERRKGVQSSGIMLTFWLVLCALAILRSKIMTALKEDAQVDFRDITFYVYFS 180

QY 181 LLLIQLVLSGSDRPLSETHDNPCESSASFLSRITFWITGLIVRGYRQLEGSD 240  
 DB 181 LLLIQLVLSGSDRPLSETHDNPCESSASFLSRITFWITGLIVRGYRQLEGSD 240

QY 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPVVYSSKQPAQKESKVDANEVEAL 300  
 DB 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPVVYSSKQPAQKESKVDANEVEAL 300

QY 301 IVKSPQKWNPSLFKVLKTYTGPYFLMSFFPKAIHDLMMFSGPQLKLLIFVNDTKAPD 360  
 DB 301 IVKSPQKWNPSLFKVLKTYTGPYFLMSFFPKAIHDLMMFSGPQLKLLIFVNDTKAPD 360

QY 361 WQGYFTVLLFVTTACQLVLHXYPHICFVSGMRITKTAIVGAVYKALVITNSARKSTV 420  
 DB 361 WQGYFTVLLFVTTACQLVLHXYPHICFVSGMRITKTAIVGAVYKALVITNSARKSTV 420

QY 421 GEIVNLSVDQRFMDLATYINIMWSAPLOVILALYLLWNLGSPVLGAVVWVLPVN 480  
 DB 421 GEIVNLSVDQRFMDLATYINIMWSAPLOVILALYLLWNLGSPVLGAVVWVLPVN 480

QY 481 AVMAKTKTYQVAHMKSDNRKILMNEILNGIKVLKYAWELAFKDKVLAIROBELKVLK 540  
 DB 481 AVMAKTKTYQVAHMKSDNRKILMNEILNGIKVLKYAWELAFKDKVLAIROBELKVLK 540

QY 541 KSAYLSAVGTFTWCTPFLVALCTPAVVTVDENNILDQAQAFVSLALFNILRPLNLP 600  
 DB 541 KSAYLSAVGTFTWCTPFLVALCTPAVVTVDENNILDQAQAFVSLALFNILRPLNLP 600

QY 601 MVISIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGGNSITVRNATFTWARSDDPT 660  
 DB 601 MVISIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGGNSITVRNATFTWARSDDPT 660

QY 661 LINGTFTSPEGALVAVVQVCGGKSSLLSALLAEMDKVEGHVAKGSAVYVPOQAWTQND 720  
 DB 661 LINGTFTSPEGALVAVVQVCGGKSSLLSALLAEMDKVEGHVAKGSAVYVPOQAWTQND 720

QY 721 SLRENILFGCOLERPYRSVIOACALLPDLEILPSGDRTEIGEKNVLSGGQKORVSLAR 780  
 DB 721 SLRENILFGCOLERPYRSVIOACALLPDLEILPSGDRTEIGEKNVLSGGQKORVSLAR 780

QY 781 AVYNADIIYLFDDPLSAVDHVGXHFENVIKPGMLKNKTRILVTHSMSYLPQVDVIIV 840  
 DB 781 AVYNADIIYLFDDPLSAVDHVGXHFENVIKPGMLKNKTRILVTHSMSYLPQVDVIIV 840

QY 841 MSGGKISEMGVSOELLARDGFAFLRTYASTEOBDAENGVTGVSQPGKEAKQMGNGM 900  
 DB 841 MSGGKISEMGVSOELLARDGFAFLRTYASTEOBDAENGVTGVSQPGKEAKQMGNGM 900

QY 901 LVTDAGKQLOQLSSSSSYSGDISRHNSHTAELOKAEAKKETWKLMEADKAQTQGVKL 960  
 DB 901 LVTDAGKQLOQLSSSSSYSGDISRHNSHTAELOKAEAKKETWKLMEADKAQTQGVKL 960

QY 961 SVYDYMKAIGLFIISFLIFCMCNHVSALASNYWLSLWTDPIVNGTQEHKTVRLSVYG 1020  
 DB 961 SVYDYMKAIGLFIISFLIFCMCNHVSALASNYWLSLWTDPIVNGTQEHKTVRLSVYG 1020

QY 1021 ALGISQGIAGFCYSMAVISIGGILASRCILHVDLLHSILRSPMSFFERTSGNLVNRFSKEL 1080  
 DB 1021 ALGISQGIAGFCYSMAVISIGGILASRCILHVDLLHSILRSPMSFFERTSGNLVNRFSKEL 1080

QY 1081 DTVDMSIPEVIXMPFMSLFPNVIGACIVILLATPIAAIIIPPLGLIYFFVQRPVYASSROL 1140  
 DB 1081 DTVDMSIPEVIXMPFMSLFPNVIGACIVILLATPIAAIIIPPLGLIYFFVQRPVYASSROL 1140

QY 1141 KRLESVSRSPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENQKAYYPSIVANRWLA 1200  
 DB 1141 KRLESVSRSPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENQKAYYPSIVANRWLA 1200

QY 1201 VRLECVGNCIVLFAALFAVISHRSLSAGLVSVSYSLQVTTYLNLVVRMSSEMETNIVA 1260  
 DB 1201 VRLECVGNCIVLFAALFAVISHRSLSAGLVSVSYSLQVTTYLNLVVRMSSEMETNIVA 1260

QY 1261 VERLKEYSETKEAPWQIQTETAPPSSWPQGRVFRNRYCLRYREDLDFVLRHINVTINGG 1320  
 DB 1261 VERLKEYSETKEAPWQIQTETAPPSSWPQGRVFRNRYCLRYREDLDFVLRHINVTINGG 1320

QY 1321 EKVGIIVGRTGAGKSLTLGLFRINESAGEEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
 DB 1321 EKVGIIVGRTGAGKSLTLGLFRINESAGEEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380

QY 1381 SGSLRMNLDPFQSYSDSEVWTSLELAHLKDFVSALPKLDHECAEGGNSLVGQRLVCL 1440  
 DB 1381 SGSLRMNLDPFQSYSDSEVWTSLELAHLKDFVSALPKLDHECAEGGNSLVGQRLVCL 1440

QY 1441 ARALLRKTILVLEDEATAVDLETTDLTQSTIRTFQFEDCTVLTIAHRLNTIMDYTRVIL 1500  
 DB 1441 ARALLRKTILVLEDEATAVDLETTDLTQSTIRTFQFEDCTVLTIAHRLNTIMDYTRVIL 1500

QY 1501 DKGBIQEYCAPSDLLQQRGLFYSMAXDAGLV 1531  
 DB 1501 DKGBIQEYCAPSDLLQQRGLFYSMAXDAGLV 1531

RESULT 2  
 AAW74471  
 ID AAW74471 standard; protein, 1531 AA.  
 XX  
 AC AAW74471;  
 XX  
 DT 18-MAY-1999 (first entry)  
 XX  
 DE Human multidrug resistance-associated protein variant.  
 XX  
 KW Multidrug resistance-associated protein; MDR; human; diagnosis;  
 KW MDR tumour cell identification; cancer therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US5882875-A.  
 XX  
 PD 16-MAR-1999.  
 XX  
 PF 05-JUN-1995; 95US-00462109.  
 XX  
 PR 27-OCT-1992; 92US-00966923.  
 PR 08-MAR-1993; 93US-00029340.  
 PR 26-OCT-1993; 93US-00141893.  
 PR 20-MAR-1995; 95US-00407207.  
 XX  
 PA (TOOH ) UNIV QUEENS KINGSTON.

XX Cole SPC, Deeley RG;  
 XX WPI: 1999-214061/18.  
 XX N-PSDB; AAX21977.  
 XX Identifying a multidrug resistant tumor cell by contacting the cell with  
 PT an antibody/antigen-binding fragment - which binds to an expressed  
 PT protein encoded by multidrug resistance-associated protein (MRP) nucleic  
 PT acid.  
 XX  
 XX Claim 3; Col 69-80; 80pp; English.  
 XX  
 CC This sequence is the human multidrug resistance-associated (MDR) protein.  
 CC The invention relates to a method for identifying a multidrug resistant  
 CC (MDR) tumour cell. Compositions and methods utilising the MDR proteins  
 CC can be used to treat patients with tumours displaying multidrug  
 CC resistance, particularly those displaying resistance to anthracyclines,  
 CC epipodophylotoxins, vinca alkaloids, and hydrophobic drugs. The methods  
 CC for inhibiting/killing a MDR tumour cell can be useful for treating  
 CC breast cancer, leukaemias, fibrosarcomas, cervical cancer, gliomas,  
 CC thymas, neuroblastomas and lung cancer. The MDR DNA sequences when  
 CC labeled are useful as molecular probes for diagnosing multidrug  
 CC resistance of a tumour (using cells from a tumour biopsy) and for  
 CC designing ribozymes which are capable of cleaving a single-stranded  
 CC nucleic acid encoding a protein having MRP activity. Recombinant  
 CC expression vectors containing human MDR coding sequences can be  
 CC transfected into a drug sensitive cell line to produce a protein in the  
 CC cell which confers MDR, protecting non-resistant non-tumour cells from  
 CC the effects of chemotherapeutics has major clinical importance. Cells  
 CC transformed with the MDR coding sequences are useful for testing  
 CC potential therapeutic agents for their effectiveness against MDR cells  
 CC and for identifying chemosensitisers of a therapeutic agent  
 XX  
 XX Sequence 1531 AA;  
 SQ  
 Query Match 78.5%; Score 7860; DB 2; Length 1531;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MALRGCSDGSDPLDWDNVTWNTSNPDTKCFQNTVWVPCFYLWACFPFYLYLSRH 60  
 DB 1 MALRGCSDGSDPLDWDNVTWNTSNPDTKCFQNTVWVPCFYLWACFPFYLYLSRH 60  
 QY 61 DRGYTQMTPLNKTALGFLMIVCWADLFYFWSRSGIFLAPVFLVSPDLLGTTLLA 120  
 DB 61 DRGYTQMTPLNKTALGFLMIVCWADLFYFWSRSGIFLAPVFLVSPDLLGTTLLA 120  
 QY 121 TFLIQLERRKGVOSGIMLTFWLVALCALAILRSKIMTALKEDAQVDFRDITFYVYFS 180  
 DB 121 TFLIQLERRKGVOSGIMLTFWLVALCALAILRSKIMTALKEDAQVDFRDITFYVYFS 180  
 QY 181 LLLIQLVLSCFDRSPFSETTHDNPCESSASFLSRITFWITGLIVRGYRQPLEGSD 240  
 DB 181 LLLIQLVLSCFDRSPFSETTHDNPCESSASFLSRITFWITGLIVRGYRQPLEGSD 240  
 QY 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPVVYSSKDPAPKESKVDANEVEAL 300  
 DB 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPVVYSSKDPAPKESKVDANEVEAL 300  
 QY 301 IVKSPQKWNPSLFKVLKTYTGFYFLMSFFFKAIHDLMMFSGPQILKLIKFVNDTKAPD 360  
 DB 301 IVKSPQKWNPSLFKVLKTYTGFYFLMSFFFKAIHDLMMFSGPQILKLIKFVNDTKAPD 360  
 QY 361 WQGYFTVLLFTTACLOTLLVHQYFHCIFVSGMRKTAIVGAYRKALVITNSARKSSTV 420  
 DB 361 WQGYFTVLLFTTACLOTLLVHQYFHCIFVSGMRKTAIVGAYRKALVITNSARKSSTV 420  
 QY 421 GEIVNLSVDAORFMDLATYINNIWSAPLOVILALYLLNLPSPVLAGVAVVWLMVFN 480  
 DB 421 GEIVNLSVDAORFMDLATYINNIWSAPLOVILALYLLNLPSPVLAGVAVVWLMVFN 480  
 QY 481 AVAMKTKTYQVAHMKSKDNRIKLAENILNGIKVLKLYAWELAFKDKVLAIROBELKVLK 540

481 AVAMKTKTYQVAHMKSKDNRIKLAENILNGIKVLKLYAWELAFKDKVLAIROBELKVLK 540  
 QY 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTDENNILDAQTAFVSLALFNILRFPNLNIP 600  
 DB 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTDENNILDAQTAFVSLALFNILRFPNLNIP 600  
 QY 601 MVISSIVQASVSLKELRIFLSHEELEPDSIERRPVKDGCGGTSITVRNATFTWARSDDPT 660  
 DB 601 MVISSIVQASVSLKELRIFLSHEELEPDSIERRPVKDGCGGTSITVRNATFTWARSDDPT 660  
 QY 661 LMGITFSIPEGALVAVVQVCGGKSSLLSALLAEMDKVEGHVAIKGSVAVYVQQAWIQND 720  
 DB 661 LMGITFSIPEGALVAVVQVCGGKSSLLSALLAEMDKVEGHVAIKGSVAVYVQQAWIQND 720  
 QY 721 SURENILFGCOLLEPYYSVIOACALLPDLETLPSGDRTEIGEGKGNVLSGGQKQVSLAR 780  
 DB 721 SURENILFGCOLLEPYYSVIOACALLPDLETLPSGDRTEIGEGKGNVLSGGQKQVSLAR 780  
 QY 781 AVYSNADIYLFDDPLSAVDHVGKHIFENVIGPKGMLKNKTRILVTHSMSYLPQVDVIIV 840  
 DB 781 AVYSNADIYLFDDPLSAVDHVGKHIFENVIGPKGMLKNKTRILVTHSMSYLPQVDVIIV 840  
 QY 841 MSGGKISEMSYQELLARDGAPAEPLRTYASTEQDQDAEENGVTGVSFGPKAKOMENGM 900  
 DB 841 MSGGKISEMSYQELLARDGAPAEPLRTYASTEQDQDAEENGVTGVSFGPKAKOMENGM 900  
 QY 901 LVTDSSAGKOLQORQLSSSSSYSGDISRHNSHTAELOKAEAKKEETWKLMEADKAQGVK 960  
 DB 901 LVTDSSAGKOLQORQLSSSSSYSGDISRHNSHTAELOKAEAKKEETWKLMEADKAQGVK 960  
 QY 961 SVYDYMKAIGLIFISFLSIFLFCMCHNVHSALASNYLSLWTDDDPIVNGTQHTKVRLSVYG 1020  
 DB 961 SVYDYMKAIGLIFISFLSIFLFCMCHNVHSALASNYLSLWTDDDPIVNGTQHTKVRLSVYG 1020  
 QY 1021 ALGISQGIATVFGYSMAVSIIGGILASRCCLHVDLLHSILRSPMSFFERTSGNVLNRPSEL 1080  
 DB 1021 ALGISQGIATVFGYSMAVSIIGGILASRCCLHVDLLHSILRSPMSFFERTSGNVLNRPSEL 1080  
 QY 1081 DTVDSMIEVIMFMGSLFNIVIGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQL 1140  
 DB 1081 DTVDSMIEVIMFMGSLFNIVIGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQL 1140  
 QY 1141 KRLESVSRSPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENQKAYYPSIVANRWLA 1200  
 DB 1141 KRLESVSRSPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENQKAYYPSIVANRWLA 1200  
 QY 1201 VRLECVGNICVILFAALPAVISHRSLSAGLVGLSVSYLSQVTTYLNLWLRMSSEMETNIVA 1260  
 DB 1201 VRLECVGNICVILFAALPAVISHRSLSAGLVGLSVSYLSQVTTYLNLWLRMSSEMETNIVA 1260  
 QY 1261 VERLKEYSETKEAPWQIQETAPSSWPQVGRVFNRYCLRYREDLDVFLRHINVTINGG 1320  
 DB 1261 VERLKEYSETKEAPWQIQETAPSSWPQVGRVFNRYCLRYREDLDVFLRHINVTINGG 1320  
 QY 1321 EKVGVIGRTGAGKSSLTGLFRINESAGEEIIIDGINIAKIGHDLRKFKITIIPQDPVLF 1380  
 DB 1321 EKVGVIGRTGAGKSSLTGLFRINESAGEEIIIDGINIAKIGHDLRKFKITIIPQDPVLF 1380  
 QY 1381 SGSLRNNLDPPSQYSDERBWTSLSLAHLKDFVSALPDKLDHECAGGENLVSGORQLVCL 1440  
 DB 1381 SGSLRNNLDPPSQYSDERBWTSLSLAHLKDFVSALPDKLDHECAGGENLVSGORQLVCL 1440  
 QY 1441 ARALRKTILVLDATAVLDLTDLLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 DB 1441 ARALRKTILVLDATAVLDLTDLLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 QY 1501 DKGEIQEYCAPSDLLQORGLFYSMADAGLV 1531  
 DB 1501 DKGEIQEYCAPSDLLQORGLFYSMADAGLV 1531

AAW99894  
 ID AAW99894 standard; protein; 1531 AA.  
 XX  
 AC AAW99894;  
 XX  
 DT 10-JUN-1999 (first entry)  
 XX  
 DE Human multidrug resistance-associated protein natural variant.  
 XX  
 KW Human; multidrug resistance-associated protein; MRP; cytotoxic drug;  
 KW cancer; chemotherapy.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US5891724-A.  
 XX  
 PD 06-APR-1999.  
 XX  
 PF 05-JUN-1995; 95US-00460907.  
 XX  
 PR 27-OCT-1992; 92US-00966923.  
 PR 08-MAR-1993; 93US-00029340.  
 PR 26-OCT-1993; 93US-00141893.  
 PR 20-MAR-1995; 95US-00407207.  
 XX  
 PA (TOOH ) UNIV QUEENS KINGSTON.  
 XX  
 PI Cole SPC, Deeley RG;  
 XX  
 DR WPI; 1999-253868/21.  
 DR N-PSDB; AAX19818.  
 XX  
 PT Protecting mammalian cells against cytotoxic drugs.  
 XX  
 PS Claim 2; Col 79-86; 82pp; English.  
 XX  
 CC The present sequence represents a human multidrug resistance-associated  
 CC protein (MRP). The present invention also describes a method for  
 CC protecting a mammalian cell against the cytotoxicity of anthracyclines,  
 CC epipodophyllotoxins and Vinca alkaloids (A) by introducing into it a  
 CC nucleic acid (I) that hybridizes under stringent conditions to a nucleic  
 CC acid (II) that encodes an MRP protein (III). Introduction of (I) protects  
 CC cells against cytotoxic effects of (A), particularly to protect normal  
 CC cells against (A) being used for treatment of cancers. Cells transformed  
 CC with (I) can be used to screen for agents that affect multidrug  
 CC resistance or are directly toxic to multidrug resistant cells, i.e.  
 CC potential therapeutics for multidrug-resistant cancers. Confering  
 CC resistance to normal cells should allow an increase in the dose of (A)  
 CC that can be administered safely  
 XX  
 SQ Sequence 1531 AA;  
 Query Match 78.5%; Score 7860; DB 2; Length 1531;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MALRGFCADGSDPLWNNWNTSNPDTKCFQNTLVWVPCFYLWACFPFFYLRSR 60  
 DB 1 MALRGFCADGSDPLWNNWNTSNPDTKCFQNTLVWVPCFYLWACFPFFYLRSR 60  
 QY 61 DRGYIQTMTPLNKTALGFLWIVCWADLFYSFWRSGIFLAPVFLVSPFLGITTLLA 120  
 DB 61 DRGYIQTMTPLNKTALGFLWIVCWADLFYSFWRSGIFLAPVFLVSPFLGITTLLA 120  
 QY 121 TFLIQLERRKGVQSSGIMLTFWLVALVCALAILRSKIMTALKEDAQVDLPDIFFYVFS 180  
 DB 121 TFLIQLERRKGVQSSGIMLTFWLVALVCALAILRSKIMTALKEDAQVDLPDIFFYVFS 180  
 QY 181 LLLIQLVLSCFSDRSPLFSETHDPNCPSSASFLSRITFTWTTIGLIVRGYRPLEGSD 240  
 DB 181 LLLIQLVLSCFSDRSPLFSETHDPNCPSSASFLSRITFTWTTIGLIVRGYRPLEGSD 240  
 QY 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKVYSSKDPAPKRESSKVDANEEVEAL 300

DB 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKVYSSKDPAPKRESSKVDANEEVEAL 300  
 QY 301 IVKSPQKEWNPFLFKVLYKTFPGYFLMGFFPKAIHDLMMFSGPQILKLLIKFVNNTKADP 360  
 DB 301 IVKSPQKEWNPFLFKVLYKTFPGYFLMGFFPKAIHDLMMFSGPQILKLLIKFVNNTKADP 360  
 QY 361 WGYFYTVLLFVTTACLOTLVLFHQYFHCIFVSGMRITKTAIVGAVYRKALVITNSARKSSTV 420  
 DB 361 WGYFYTVLLFVTTACLOTLVLFHQYFHCIFVSGMRITKTAIVGAVYRKALVITNSARKSSTV 420  
 QY 421 GEIVNLSVDAQRFMDLATYINMIWSAPLOVILALYLWNLGSPVLGAVVMLMVPVN 480  
 DB 421 GEIVNLSVDAQRFMDLATYINMIWSAPLOVILALYLWNLGSPVLGAVVMLMVPVN 480  
 QY 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAWELAFKDKVLAIROBELKVLK 540  
 DB 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAWELAFKDKVLAIROBELKVLK 540  
 QY 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTTDENNILDQAOTAFVSLALFNILRFPPLTILP 600  
 DB 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTTDENNILDQAOTAFVSLALFNILRFPPLTILP 600  
 QY 601 MVISSIVOASVSLKRLRIFLSHELEPDSIERRPVKDGGTNSITVRNATFTWARSDDPT 660  
 DB 601 MVISSIVOASVSLKRLRIFLSHELEPDSIERRPVKDGGTNSITVRNATFTWARSDDPT 660  
 QY 661 LINGITFSIPEGALVAVVGVQCGKSSLLSALLAEMDKVEGHVAIKGSVAVYVPPQAWIQND 720  
 DB 661 LINGITFSIPEGALVAVVGVQCGKSSLLSALLAEMDKVEGHVAIKGSVAVYVPPQAWIQND 720  
 QY 721 SLRENILFGCOLEBPYRSVIOACALLPDLBILPSGDRTEIGEGKVNLSGQOKORVSLAR 780  
 DB 721 SLRENILFGCOLEBPYRSVIOACALLPDLBILPSGDRTEIGEGKVNLSGQOKORVSLAR 780  
 QY 781 AVYSNADIYLFDDPLSAVDAHVGHKIPENVIGPKGMLKNKRIILVTHSMSYLPQVDVILV 840  
 DB 781 AVYSNADIYLFDDPLSAVDAHVGHKIPENVIGPKGMLKNKRIILVTHSMSYLPQVDVILV 840  
 QY 841 MSGKISSEMGSYQELLARDGAFBFLTYASTEQDAENGVTGVSQPGKEAKQEMNGM 900  
 DB 841 MSGKISSEMGSYQELLARDGAFBFLTYASTEQDAENGVTGVSQPGKEAKQEMNGM 900  
 QY 901 LVTDASGKQORQSSSSSYSGDISRHNSHSTAELOKAEKEETWKLMEADKQGTQVKL 960  
 DB 901 LVTDASGKQORQSSSSSYSGDISRHNSHSTAELOKAEKEETWKLMEADKQGTQVKL 960  
 QY 961 SVYWDYMKAIGLFISFLSIFLFCNHNVSALASNYLSLWTDPIVNGTQEHKTVRLSVYG 1020  
 DB 961 SVYWDYMKAIGLFISFLSIFLFCNHNVSALASNYLSLWTDPIVNGTQEHKTVRLSVYG 1020  
 QY 1021 ALGISQGIAPVGYSMVAVSIGILLASRCLHVDLHLSILRSPMSFFERTPSGNLVRNFSKEL 1080  
 DB 1021 ALGISQGIAPVGYSMVAVSIGILLASRCLHVDLHLSILRSPMSFFERTPSGNLVRNFSKEL 1080  
 QY 1081 DTVDSDMIPVYKFMFMSLNFVIGACIVILLATPAAIIIPPLGLIYFVFORFYVASSRQL 1140  
 DB 1081 DTVDSDMIPVYKFMFMSLNFVIGACIVILLATPAAIIIPPLGLIYFVFORFYVASSRQL 1140  
 QY 1141 KRLESVSRSPVYSHFNETLLGVSVIRAFEPQERFIHQSDLKVDENQKAYYPSIVANRWLA 1200  
 DB 1141 KRLESVSRSPVYSHFNETLLGVSVIRAFEPQERFIHQSDLKVDENQKAYYPSIVANRWLA 1200  
 QY 1201 VRLECVCNCILVFAALFAVLSRHSLSAGLVLSVSYSLQVTTYLNLVLRMSSEMETNIVA 1260  
 DB 1201 VRLECVCNCILVFAALFAVLSRHSLSAGLVLSVSYSLQVTTYLNLVLRMSSEMETNIVA 1260  
 QY 1261 VERLKEYSETEKAPMQIOETAPPPSWPQVGRVFEFNYCLIRYREDLDFVLRHINVTINGG 1320  
 DB 1261 VERLKEYSETEKAPMQIOETAPPPSWPQVGRVFEFNYCLIRYREDLDFVLRHINVTINGG 1320  
 QY 1321 EKVGIVGRTGAGKSSITGLGFRINESABGEIIGDGINIAKIGLHDLRFKTIIPQDPVLV 1380

Db 1321 EKVGIVGRTGAGKSLTJGLFRINSAEGBEIIIDGINIAKIGHDLRKFKITIIPQDPVLF 1380  
 QY 1381 SGLSRNLNDPPFQYSDDEEVTSLAHLKDFVSALPDKLDHCAEGGNSLVGQRLVCL 1440  
 Db 1381 SGLSRNLNDPPFQYSDDEEVTSLAHLKDFVSALPDKLDHCAEGGNSLVGQRLVCL 1440  
 QY 1441 ABALLRKTKILVLDATAAVDLETDLLIQSTTRTQPEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 Db 1441 ABALLRKTKILVLDATAAVDLETDLLIQSTTRTQPEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 QY 1501 DKGEIOEYCAPSDLLQOQGLFYSMKADAGLV 1531  
 Db 1501 DKGEIOEYCAPSDLLQOQGLFYSMKADAGLV 1531

RESULT 4

ID AAY55799 standard; protein; 1531 AA.  
 XX AAY55799;  
 XX 28-FEB-2000 (first entry)  
 XX Human multidrug resistance-associated protein (MRP) variant.  
 XX Chemosensitizer; multidrug resistance-associated protein; MRP; human;  
 KW therapeutic agent; P-glycoprotein-mediated multidrug resistance; lung;  
 KW cancer; variant.  
 XX Homo sapiens.

Key Location/Qualifiers  
 FT Misc-difference 685 /label= L68SS  
 FT /note= "wild-type Leu is replaced by Ser"  
 FT Misc-difference 1282 /label= R1282A  
 FT /note= "wild-type Arg is replaced by Ala"

US6001563-A.  
 PD 14-DEC-1999.  
 XX 05-JUN-1995; 95US-00463179.  
 XX 27-OCT-1992; 92US-00966923.  
 PR 08-MAR-1993; 93US-00029340.  
 PR 26-OCT-1993; 93US-00141893.  
 PR 20-MAR-1995; 95US-00407207.  
 XX (TOOH ) UNIV QUEENS KINGSTON.  
 XX Cole SP, Deeley RG;  
 XX WPI; 2000-061877/05.  
 DR N-PSDB; AA239556.  
 XX Identification of chemosensitizers useful for treating cancer, using  
 PT nucleic acids encoding multidrug resistance-associated protein.  
 XX Claim 3; Col 69-80; 77pp; English.  
 XX The invention provides a method for identifying a substance which is a  
 CC chemosensitizer that comprises, contacting a cell transfected with  
 CC nucleic acid encoding multidrug resistance-associated protein (MRP) with  
 CC a therapeutic agent in vitro. The method is useful for identifying  
 CC chemosensitizers which may then be used to treat cancer (especially lung  
 CC cancer). The method allows the identification of chemosensitizers which  
 CC do not reverse P-glycoprotein-mediated multidrug resistance. The present  
 CC sequence represents a human MRP variant  
 XX Sequence 1531 AA;

Query Match	78.5%;	Score 7860;	DB 3;	Length 1531;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1531;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

  

QY	1	MALRGFCSDGSDPLMDWNNVTWNTSNPDFTKCFQNTVLVWVPCFVLMACFPFYLYLSH	60
Db	1	MALRGFCSDGSDPLMDWNNVTWNTSNPDFTKCFQNTVLVWVPCFVLMACFPFYLYLSH	60
QY	61	DRGYQMTPLNKTALGFLWVWADLFLYSFWSRSRGIFLAPVFLVSPDLLGITLLA	120
Db	61	DRGYQMTPLNKTALGFLWVWADLFLYSFWSRSRGIFLAPVFLVSPDLLGITLLA	120
QY	121	TFLIOLERRKGVQSSGIMLTFWLVALCALAILRSKIMTALKEDAQVDLFRDITFVYFS	180
Db	121	TFLIOLERRKGVQSSGIMLTFWLVALCALAILRSKIMTALKEDAQVDLFRDITFVYFS	180
QY	181	LLLIQLVLSCSRSPLESETHDNPCESSASFLSRITFWITGLIVRGVROPLEGSD	240
Db	181	LLLIQLVLSCSRSPLESETHDNPCESSASFLSRITFWITGLIVRGVROPLEGSD	240
QY	241	LWSLNKEDTSEOVVPLVKNWKECAKTRKQPVKVYVSSKDPAPQKSSKVDANEVEAL	300
Db	241	LWSLNKEDTSEOVVPLVKNWKECAKTRKQPVKVYVSSKDPAPQKSSKVDANEVEAL	300
QY	301	IVKSPQKWNPSLFKLYKTGYPFLMSFFFKAIHDLMMFSGPQILKLLIKFVNDTKAPD	360
Db	301	IVKSPQKWNPSLFKLYKTGYPFLMSFFFKAIHDLMMFSGPQILKLLIKFVNDTKAPD	360
QY	361	WQGYFTVLLFVTAQLTLVHLYPHICFVSGMRITKAVIGAVYRKALVITNSARKSSTV	420
Db	361	WQGYFTVLLFVTAQLTLVHLYPHICFVSGMRITKAVIGAVYRKALVITNSARKSSTV	420
QY	421	GEIVNLSVDAQRFMDLATYINMWSAPLOVTLALYLLWNLGSPSLAGVAVMLMVPVN	480
Db	421	GEIVNLSVDAQRFMDLATYINMWSAPLOVTLALYLLWNLGSPSLAGVAVMLMVPVN	480
QY	481	AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKLYAWELAFKDKVLAIROBELKVLK	540
Db	481	AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKLYAWELAFKDKVLAIROBELKVLK	540
QY	541	KSAYLSAVGTTFWCTPFLVALCTFAVVTVDENNILDAQTAFVLSALFNILRFPNLILP	600
Db	541	KSAYLSAVGTTFWCTPFLVALCTFAVVTVDENNILDAQTAFVLSALFNILRFPNLILP	600
QY	601	MVSIIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWASDPPT	660
Db	601	MVSIIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWASDPPT	660
QY	661	LNGITFSIPEGALVAVVQVCGKSSLLSALLAEMDKVEGHVAIKGSVAVVPQAWIQND	720
Db	661	LNGITFSIPEGALVAVVQVCGKSSLLSALLAEMDKVEGHVAIKGSVAVVPQAWIQND	720
QY	721	SIRENIFGCQLEEPYRSVITQACALLPDLETLPSGDRTEICEKGVNLGGOKQVSLAR	780
Db	721	SIRENIFGCQLEEPYRSVITQACALLPDLETLPSGDRTEICEKGVNLGGOKQVSLAR	780
QY	781	AVYSNADITYLFDPLSADVAHVGHKIFENVIGPKMGLKNKTRILVTHSHSYLPQVDVIV	840
Db	781	AVYSNADITYLFDPLSADVAHVGHKIFENVIGPKMGLKNKTRILVTHSHSYLPQVDVIV	840
QY	841	MSGKISMGYSQVQELLARDGAFELRTYASTEQEODAEENGVTGVSFGKAKQWENGM	900
Db	841	MSGKISMGYSQVQELLARDGAFELRTYASTEQEODAEENGVTGVSFGKAKQWENGM	900
QY	901	LVTDSAGLQORQLSSSSSYSGDISRHNSIAELQAKAKKEETWKLMEADKAQGVKL	960
Db	901	LVTDSAGLQORQLSSSSSYSGDISRHNSIAELQAKAKKEETWKLMEADKAQGVKL	960
QY	961	SVYDYMKAIGLFTSIFLIFMCMNVHSALASNYWLSLWTDPIVNGTQEHKTVRLSVYG	1020
Db	961	SVYDYMKAIGLFTSIFLIFMCMNVHSALASNYWLSLWTDPIVNGTQEHKTVRLSVYG	1020
QY	1021	ALGISQGIATVFGYSMAVSIIGGILASRCULHVDLLHSILRSPMSFFERTPSGNLVNRFKSL	1080

1021 ALGISQIAVFGYSMAVSGIGLILASCLHVDLHLSILRPMSEFFETPSGNLVNRSEKL 1080  
 1081 DTVDSEIPEVIMKFMGSLFNVIACIVILLATPIAAIIIPPLGLIYFFVQRFVASSRQL 1140  
 1081 DTVDSEIPEVIMKFMGSLFNVIACIVILLATPIAAIIIPPLGLIYFFVQRFVASSRQL 1140  
 1141 KRLESVSRSPVYSHNETLLGVSIVIRAFEEQERFIHQSDLKVDENKAYPSIVANRWLA 1200  
 1141 KRLESVSRSPVYSHNETLLGVSIVIRAFEEQERFIHQSDLKVDENKAYPSIVANRWLA 1200  
 1201 VRLECVGNCIVLFAALFAVISRHSLSAGLVGVSYSLOVTTYLNWLVRMSSEMETNIVA 1260  
 1201 VRLECVGNCIVLFAALFAVISRHSLSAGLVGVSYSLOVTTYLNWLVRMSSEMETNIVA 1260  
 1261 VERLKEYSTEKEAPNQIQTETAPPSWPQVGRVFRNRYCLRVREDLDFVLRHINTVINGG 1320  
 1261 VERLKEYSTEKEAPNQIQTETAPPSWPQVGRVFRNRYCLRVREDLDFVLRHINTVINGG 1320  
 1321 EKVGIIVGRTGAGKSSLTGLFRINSAEGEIIIDGINIAKIGLHDLRFKTIIPDDPVLF 1380  
 1321 EKVGIIVGRTGAGKSSLTGLFRINSAEGEIIIDGINIAKIGLHDLRFKTIIPDDPVLF 1380  
 1381 SGSLRNLDPPFQYSDEEVWTSLELAHLKDFVSALPDKLDHECAGGENLSVGQRLVCL 1440  
 1381 SGSLRNLDPPFQYSDEEVWTSLELAHLKDFVSALPDKLDHECAGGENLSVGQRLVCL 1440  
 1441 ARALARKTKILVLDREATAVDLETDLIQSTIRTOFEDCTVLTIAHRLANTIMDYTRVIVL 1500  
 1441 ARALARKTKILVLDREATAVDLETDLIQSTIRTOFEDCTVLTIAHRLANTIMDYTRVIVL 1500  
 1501 DKGEIQEYGAPELDLQORGLFYSMADAGLV 1531  
 1501 DKGEIQEYGAPELDLQORGLFYSMADAGLV 1531

## RESULT 5

AAY78873  
 ID AAY78873 standard; protein; 1531 AA.

AC AAY78873;

DT 19-MAY-2000 (first entry)

DE Multidrug resistance protein (MRP) natural variant amino acid sequence.

KW Multidrug resistance protein; MRP; human; anthracycline; vinca alkaloid;  
 KW epipodophyllotoxin; cancer; leukaemia.

XX Homo sapiens.

OS US6025473-A.

PN 15-FEB-2000.

XX 05-JUN-1995; 95US-00461384.

XX 27-OCT-1992; 92US-00966923.

XX 08-MAR-1993; 93US-00029340.

XX 26-OCT-1993; 93US-00141893.

XX 20-MAR-1995; 95US-00407207.

XX (TOOH ) UNIV QUEEN'S KINGSTON.

XX Cole SPC, Deeley RG;

XX WPI; 2000-181838/16.

XX N-PSDB; AAZ90193.

PT Isolated protein conferring multidrug resistance, to at least two drugs  
 PT selected from anthracyclines, epipodophyllotoxins and vinca alkaloids, on  
 PT a drug sensitive mammalian cell.

XX

PS Claim 10; Col 79-88; 78pp; English.

XX This sequence represents a human multidrug resistance protein (MRP)  
 CC natural variant amino acid sequence. The human MRP confers multidrug  
 CC resistance, including resistance to at least two drugs selected from  
 CC anthracyclines, epipodophyllotoxins and vinca alkaloids, on a drug  
 CC sensitive mammalian cell, when the protein is expressed in the cell. The  
 CC multidrug resistance is not substantially reversed by chemosensitizers  
 CC which reverse P-glycoprotein-mediated multidrug resistance. The MRP  
 CC protein sequence can be used to generate antibodies against MRP. The MRP  
 CC protein and nucleotide sequences can be used in compositions which are  
 CC used to treat patients with tumours displaying multidrug resistance. The  
 CC compositions and methods of the invention can be used particularly to  
 CC treat breast cancer, leukaemias, fibrosarcomas, cervical cancer, gliomas,  
 CC thymomas, neuroblastomas, and lung cancer. Antibodies directed against  
 CC MRP can be used to inhibit the multidrug resistance of a multidrug  
 CC resistant cell

SQ Sequence 1531 AA;

Query Match 78.5%; Score 7860; DB 3; Length 1531;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCSDGSDPLDNDNVTWNTSNPDFTKCFQNTVLVWVPCFYLMACFPFFYLYLSRH 60  
 DB 1 MALRGFCSDGSDPLDNDNVTWNTSNPDFTKCFQNTVLVWVPCFYLMACFPFFYLYLSRH 60  
 QY 61 DRGIQMTPLNKTALGFLIIVCWADLFYSFERSRGIFLAPVFLVSPFLGTTLLA 120  
 DB 61 DRGIQMTPLNKTALGFLIIVCWADLFYSFERSRGIFLAPVFLVSPFLGTTLLA 120  
 QY 121 TFLIQLERRKGVQSSGIMLTFWLVALCALAILRSKIMTALKEDAQVDLFRDITFYVVS 180  
 DB 121 TFLIQLERRKGVQSSGIMLTFWLVALCALAILRSKIMTALKEDAQVDLFRDITFYVVS 180  
 QY 181 LLLQLVLSCFSDRSPFLFSETHDNPSPSSASFLSRTTFWMTGLIVRGYRPLSGSD 240  
 DB 181 LLLQLVLSCFSDRSPFLFSETHDNPSPSSASFLSRTTFWMTGLIVRGYRPLSGSD 240  
 QY 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKVYSSKDPAPQKSSKVDANEEVEAL 300  
 DB 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKVYSSKDPAPQKSSKVDANEEVEAL 300  
 QY 301 IVKSPQKEWNPFLFKLYKTGPGYFLMSFFPKAHDLMFSGPOILKLLIKFVNDTKAPD 360  
 DB 301 IVKSPQKEWNPFLFKLYKTGPGYFLMSFFPKAHDLMFSGPOILKLLIKFVNDTKAPD 360  
 QY 361 WQGYFTVLLFVTAQLTLVHLQVFIHCFVSGMRITKTAIVGAVYRKALVITNSARKSSTV 420  
 DB 361 WQGYFTVLLFVTAQLTLVHLQVFIHCFVSGMRITKTAIVGAVYRKALVITNSARKSSTV 420  
 QY 421 GEIVNLSVDAQRFMDLATYINNIWSAPLOVILALYLLMLNLGFSVLGAVVMVLPVN 480  
 DB 421 GEIVNLSVDAQRFMDLATYINNIWSAPLOVILALYLLMLNLGFSVLGAVVMVLPVN 480  
 QY 481 AVNAMKTKYQVAHMSKDNRIKLMNEILNGIKVLKYAWELAFKDKVLAIROEELKVLK 540  
 DB 481 AVNAMKTKYQVAHMSKDNRIKLMNEILNGIKVLKYAWELAFKDKVLAIROEELKVLK 540  
 QY 541 KSAVLSAVGTFTWVCTPFLVALCTFAVYVITDENNILDATQAFVSLALFNILFPNLP 600  
 DB 541 KSAVLSAVGTFTWVCTPFLVALCTFAVYVITDENNILDATQAFVSLALFNILFPNLP 600  
 QY 601 MVISSIVQASVSLKRLRIFLSHEELPDSIERPPVKDGGTNSITVRNATFTWARSDDPT 660  
 DB 601 MVISSIVQASVSLKRLRIFLSHEELPDSIERPPVKDGGTNSITVRNATFTWARSDDPT 660  
 QY 661 LNCITTSIPGALVAVVGQVCGKSLLSALLAEMDKVEGHVAIKGSVAVVPQAMQND 720  
 DB 661 LNCITTSIPGALVAVVGQVCGKSLLSALLAEMDKVEGHVAIKGSVAVVPQAMQND 720  
 QY 721 SURENIFGQCEEPYRYSVIQACALLPDLLEILPSGDRTEIGEGVNLSSGQKQVSLAR 780

Db 721 SLRENILFCQLEPPYRVIQACALLPDLLEILPSGDRTEIGEGKGNLSGGQKQVSLAR 780  
 QY 781 AVYSNADIYLFDDPLSAVDHVGKHFENVIQKGMKKNKTRILVTHSMSYLPQVDVIV 840  
 Db 781 AVYSNADIYLFDDPLSAVDHVGKHFENVIQKGMKKNKTRILVTHSMSYLPQVDVIV 840  
 QY 841 MSGGKISENGSYQELLARDGAFALRTYASTEQEQDAEENGVTGSGFGKEAKQWENG 900  
 Db 841 MSGGKISENGSYQELLARDGAFALRTYASTEQEQDAEENGVTGSGFGKEAKQWENG 900  
 QY 901 LVTDSAGKQLQRLSSSSSYSGDISRHNSHSTAELOKAEAKKEETWKLMEADYAOQTQV 960  
 Db 901 LVTDSAGKQLQRLSSSSSYSGDISRHNSHSTAELOKAEAKKEETWKLMEADYAOQTQV 960  
 QY 961 SVYDYMKAIGLFIISFLSIFLFCMCHVLSALSNYLSLWTDPIVNGTOEHTKVLRSYVG 1020  
 Db 961 SVYDYMKAIGLFIISFLSIFLFCMCHVLSALSNYLSLWTDPIVNGTOEHTKVLRSYVG 1020  
 QY 1021 ALGISQGIIVFGYSMAVSTGGILASRCLHVDLLHSILRSPMFFERTPSGNLVNRFKEL 1080  
 Db 1021 ALGISQGIIVFGYSMAVSTGGILASRCLHVDLLHSILRSPMFFERTPSGNLVNRFKEL 1080  
 QY 1081 DTVDMSIPVIMKPMGSLFNVIACIVILLATPIAAIIIPPLGLIYFFVQRYVASSRQL 1140  
 Db 1081 DTVDMSIPVIMKPMGSLFNVIACIVILLATPIAAIIIPPLGLIYFFVQRYVASSRQL 1140  
 QY 1141 KLESVSRSPVYSHNETLLGVSVTRAPEEQERFIHQSDLKVDENOKAYPSIVANRWLA 1200  
 Db 1141 KLESVSRSPVYSHNETLLGVSVTRAPEEQERFIHQSDLKVDENOKAYPSIVANRWLA 1200  
 QY 1201 VRLCVCNGCIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTYLNWLRMSSEMETNIVA 1260  
 Db 1201 VRLCVCNGCIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTYLNWLRMSSEMETNIVA 1260  
 QY 1261 VERLKEYSTEAPWQIETAPPSPQVGRVFRNYCLRYREDLDFVLRHINTYINGG 1320  
 Db 1261 VERLKEYSTEAPWQIETAPPSPQVGRVFRNYCLRYREDLDFVLRHINTYINGG 1320  
 QY 1321 EVKGVIGRTGAGKSSLTGLFRINSAEGLIIDGINIAKIGLHDLRFKTIIPDPVLF 1380  
 Db 1321 EVKGVIGRTGAGKSSLTGLFRINSAEGLIIDGINIAKIGLHDLRFKTIIPDPVLF 1380  
 QY 1381 SGSLRNLDPFQYSDVEEVTLSLELAHLKDFVSALPDKLDHECAGGENLSVGQRLVCL 1440  
 Db 1381 SGSLRNLDPFQYSDVEEVTLSLELAHLKDFVSALPDKLDHECAGGENLSVGQRLVCL 1440  
 QY 1441 ABALLRKTKILVLDENATAVDLETDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 Db 1441 ABALLRKTKILVLDENATAVDLETDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 QY 1501 DKGEIQEYCAPSDDLQOQRLGFYSMAKDAGLV 1531  
 Db 1501 DKGEIQEYCAPSDDLQOQRLGFYSMAKDAGLV 1531

RESULT 6  
 ID ABG61810 standard; protein; 1531 AA.  
 XX AC ABG61810;  
 XX DT 15-AUG-2002 (first entry)  
 XX DE Prostate cancer-associated protein #11.  
 XX KW Prostate cancer; prostate tumour tissue; human; mammal; cytostatic.  
 XX OS Mammalia.  
 XX PN WO2002020268-A2.  
 XX PD 18-APR-2002.

XX 12-OCT-2001; 2001WO-US032045.  
 XX 13-OCT-2000; 2000US-00687576.  
 PR 08-DEC-2000; 2000US-00733288.  
 PR 08-DEC-2000; 2000US-00733742.  
 PR 24-JAN-2001; 2001US-0263957P.  
 PR 16-MAR-2001; 2001US-0276791P.  
 PR 16-MAR-2001; 2001US-0276889P.  
 PR 06-APR-2001; 2001US-0281922P.  
 PR 24-APR-2001; 2001US-0286214P.  
 PR 30-APR-2001; 2001US-00847046.  
 PR 04-MAY-2001; 2001US-0088589P.  
 XX (EOSB-) EOS BIOTECHNOLOGY INC.  
 XX Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;  
 WPI; 2002-471335/50.  
 XX N-PSDB; ABK92125.  
 DR Detecting a prostate cancer-associated transcript in a cell in a patient,  
 PT useful for diagnosing prostate cancer (PC) or screening modulators of PC,  
 PT by determining if prostate cancer-associated genes are expressed in a  
 PT prostate tissue.  
 PS Claim 27; Page 309; 436pp; English.  
 XX The present invention relates to methods of detecting a prostate cancer-  
 CC associated transcript in a cell from a patient. The method comprises  
 CC contacting a biological sample from the patient with prostate cancer-  
 CC associated polynucleotides (designated PC genes) that selectively  
 CC hybridise to a sequence that is at least 80% identical to them. The  
 CC prostate cancer-associated polynucleotide sequences are differentially  
 CC expressed in prostate tumour tissue or in prostate cancer and are derived  
 CC from the tissues of various organisms such as humans or other mammals  
 CC (e.g. mice, sheep and dogs). The methods of the invention are useful for  
 CC diagnosing and treating prostate cancer in mammals. The prostate cancer-  
 CC associated genes are useful for diagnosing or treating prostate cancer,  
 CC as well as for identifying modulators of prostate cancer or agents that  
 CC inhibit prostate cancer. The nucleic acid sequences are particularly  
 CC useful in gene therapy, as a vaccine or in antisense applications.  
 CC ABG61800-ABG61944 represent prostate cancer-associated proteins  
 XX Sequence 1531 AA;

Query Match 78.5%; Score 7860; DB 5; Length 1531;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MALRGFCSADGSDPLMDNVTWNTSNPDFTKCFQNTVLVWVPCFVLWACFPFYLYLSRH 60  
 Db 1 MALRGFCSADGSDPLMDNVTWNTSNPDFTKCFQNTVLVWVPCFVLWACFPFYLYLSRH 60  
 QY 61 DRGYIQTMTPLNKTALGFLWIVCWADLFYSFWEERSRGIFLAPVFLVSPITLLGTTLLA 120  
 Db 61 DRGYIQTMTPLNKTALGFLWIVCWADLFYSFWEERSRGIFLAPVFLVSPITLLGTTLLA 120  
 QY 121 TFLIQLERRKGVQSSGIMLTFWLVLCALAILRSKINTALKEDAQVDLFRDITFVYVPS 180  
 Db 121 TFLIQLERRKGVQSSGIMLTFWLVLCALAILRSKINTALKEDAQVDLFRDITFVYVPS 180  
 QY 181 LLLIQLVLSCFSDRSPLFSETHDPNCPSSASFLSRITFWITGLIVRGVQPLEGSD 240  
 Db 181 LLLIQLVLSCFSDRSPLFSETHDPNCPSSASFLSRITFWITGLIVRGVQPLEGSD 240  
 QY 241 LWSLNKEDTSBQVVPVLVKNWKKCAKTRKQPKVYVSSKDPAPQKSSKVDANEVEAL 300  
 Db 241 LWSLNKEDTSBQVVPVLVKNWKKCAKTRKQPKVYVSSKDPAPQKSSKVDANEVEAL 300  
 QY 301 IVKSPQKEWNPFLFKVLYKTGPGPYFLMSFFPKAHDLMFSGPQILKLLIKFVNDTKAPD 360  
 Db 301 IVKSPQKEWNPFLFKVLYKTGPGPYFLMSFFPKAHDLMFSGPQILKLLIKFVNDTKAPD 360

QY 361 WQGYFYTVLLFVTAQLQTLVHLQYFHCYVSGMRIRKTAIVIGAVYRKALVITNSARKSSTV 420  
 Db 361 WQGYFYTVLLFVTAQLQTLVHLQYFHCYVSGMRIRKTAIVIGAVYRKALVITNSARKSSTV 420  
 QY 421 GEIVNLMSVDAQRFMDLATVINNIWSAPLQVILALYLLMLNLGSPSLAGVAVVNLMPVN 480  
 Db 421 GEIVNLMSVDAQRFMDLATVINNIWSAPLQVILALYLLMLNLGSPSLAGVAVVNLMPVN 480  
 QY 481 AVAMKTKYQVAHMSKONRIKLNELINGIKVLYAYWELAFKDKVLAIRQBELKVLK 540  
 Db 481 AVAMKTKYQVAHMSKONRIKLNELINGIKVLYAYWELAFKDKVLAIRQBELKVLK 540  
 QY 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTDENNILDQAQFVSLALNFILRFLNINLP 600  
 Db 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTDENNILDQAQFVSLALNFILRFLNINLP 600  
 QY 601 MWISSIVQASVSLKRLRIFLSHEELEPDSIERPVKDGCGTNSITVRNATFTWASDPPT 660  
 Db 601 MWISSIVQASVSLKRLRIFLSHEELEPDSIERPVKDGCGTNSITVRNATFTWASDPPT 660  
 QY 661 LINGITFSIEGALVAVVGQVCGKSSLLSALLAEMDKVEGHVAIKGSVAYVPOQAWIQND 720  
 Db 661 LINGITFSIEGALVAVVGQVCGKSSLLSALLAEMDKVEGHVAIKGSVAYVPOQAWIQND 720  
 QY 721 SLRENILFCQLEEPYRSVIOACALLPDLILPSGDRTEIGEGKGNLSGGQKQVSLAR 780  
 Db 721 SLRENILFCQLEEPYRSVIOACALLPDLILPSGDRTEIGEGKGNLSGGQKQVSLAR 780  
 QY 781 AYSNADIYLPDPLSANDAHVGHKIHFNVIQPKMLKXKTRILVTHSMSYLPQVDVIV 840  
 Db 781 AYSNADIYLPDPLSANDAHVGHKIHFNVIQPKMLKXKTRILVTHSMSYLPQVDVIV 840  
 QY 841 MSGGKISEMGSYQELLARDGAEFLRTVASTEQDQAEENGVTGSGPGKAKOMENG 900  
 Db 841 MSGGKISEMGSYQELLARDGAEFLRTVASTEQDQAEENGVTGSGPGKAKOMENG 900  
 QY 901 LVTDSAGKOLQOLSSSSSYSGDI SRHNSSTAELQKAEAKKEETWKLMEADKAQGTQVKL 960  
 Db 901 LVTDSAGKOLQOLSSSSSYSGDI SRHNSSTAELQKAEAKKEETWKLMEADKAQGTQVKL 960  
 QY 961 SVYDYMKAIGLFIPLSIFLPMCHNVSLASNYLSLWTDPIVNGTQEHKTVLSYVG 1020  
 Db 961 SVYDYMKAIGLFIPLSIFLPMCHNVSLASNYLSLWTDPIVNGTQEHKTVLSYVG 1020  
 QY 1021 ALGISQGIAGVGSMAVSIIGTILASRCLHVLHLSILRSPMSFFERTPSGNLVRFSKEL 1080  
 Db 1021 ALGISQGIAGVGSMAVSIIGTILASRCLHVLHLSILRSPMSFFERTPSGNLVRFSKEL 1080  
 QY 1081 DTVDGMIPEVIMKFMGSLFNVTGACIVILLATPTAAIIIPPLGLIYFFVQRFYVASSRQL 1140  
 Db 1081 DTVDGMIPEVIMKFMGSLFNVTGACIVILLATPTAAIIIPPLGLIYFFVQRFYVASSRQL 1140  
 QY 1141 KLESVSRSVPVSHNETLLGVSVIRAFEEQERFTHQSDLVKDNOKAYPSIVANRWLA 1200  
 Db 1141 KLESVSRSVPVSHNETLLGVSVIRAFEEQERFTHQSDLVKDNOKAYPSIVANRWLA 1200  
 QY 1201 VRLECVGNCIVLFAALFAVIRSRHLSLGLVLSYSYSLQVTTYLNWLRMSSEMETNIVA 1260  
 Db 1201 VRLECVGNCIVLFAALFAVIRSRHLSLGLVLSYSYSLQVTTYLNWLRMSSEMETNIVA 1260  
 QY 1261 VERLKEYSETKEAPWQIQTAPSPSWPQGRVFRPNCLYRBDLPVLRSHINVTINGG 1320  
 Db 1261 VERLKEYSETKEAPWQIQTAPSPSWPQGRVFRPNCLYRBDLPVLRSHINVTINGG 1320  
 QY 1321 EKVGVGTGTGAGKSLTGLFRINESAGEEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
 Db 1321 EKVGVGTGTGAGKSLTGLFRINESAGEEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
 QY 1381 SGSLRMNLDPPSQSDEEVTWTSLELAHLKDFVSALPKDLDECAEGGENLSVGORQLVCL 1440  
 Db 1381 SGSLRMNLDPPSQSDEEVTWTSLELAHLKDFVSALPKDLDECAEGGENLSVGORQLVCL 1440

QY 1441 ARALLRKTKILVLDATRAVDLETDLLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 Db 1441 ARALLRKTKILVLDATRAVDLETDLLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 QY 1501 DKGEIQEYGAPELDLQQRGLFYSMADAGLV 1531  
 Db 1501 DKGEIQEYGAPELDLQQRGLFYSMADAGLV 1531

RESULT 7  
 ABM35012  
 ID ABM35012 standard; protein; 1531 AA.  
 XX  
 AC ABM35012;  
 XX  
 DT 08-OCT-2003 (first entry)  
 XX  
 DE Cancer based on CYP3A5 related protein SEQ ID NO:678.  
 XX  
 KW Cancer; CYP3A5; irinotecan; pharmaceutical; malignant glioma;  
 KW cytochrome p450; subfamily IIIA; nifedipine oxidase; polypeptide 5;  
 KW cytostatic.  
 XX  
 OS Unidentified.  
 XX  
 PN WO2003013534-A2.  
 XX  
 PD 20-FEB-2003.  
 XX  
 PF 23-JUL-2002; 2002WO-EP008219.  
 XX  
 PR 23-JUL-2001; 2001EP-00117608.  
 PR 24-MAY-2002; 2002EP-00011710.  
 XX  
 PA (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.  
 XX  
 PI Heinrich G, Kerb R;  
 XX  
 DR WPI; 2003-268144/26.  
 XX  
 PT New use of irinotecan for preparation of compositions for treating cancer  
 PT in subject having genome with variant allele comprising cytochrome p450,  
 PT subfamily IIIA, polypeptide 5 polynucleotide, termed CYP3A5.  
 XX  
 PS Disclosure; SEQ ID NO 678; 86pp; English.  
 XX  
 CC The present invention describes the use of irinotecan (I) or its  
 CC derivative for the preparation of a pharmaceutical composition for  
 CC treating colorectal, cervical, gastric, lung, ovarian or pancreatic  
 CC cancer, or malignant glioma in a subject having a genome with a variant  
 CC allele which comprises a cytochrome p450, subfamily IIIA (nifedipine  
 CC oxidase), polypeptide 5 (CYP3A5) polynucleotide (II). (I) and (II) have  
 CC cytostatic activity. The therapeutic applications of (I) is improved,  
 CC since it is possible to individually treat a subject with an appropriate  
 CC dosage and/or an appropriate derivative of (I). Therefore, undesirable,  
 CC harmful or toxic effects are efficiently avoided. Unnecessary and  
 CC potentially harmful treatment of those subjects who do not respond to the  
 CC treatment with substances (nonresponders), as well as the development of  
 CC drug resistances due to suboptimal drug dosing can be avoided. ACFe2200  
 CC to ACFe2751 and ABM34912 to ABM35013 represent sequences used in the  
 CC exemplification of the present invention  
 XX  
 SQ Sequence 1531 AA;

Query Match 78.5%; Score 7860; DB 6; Length 1531;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCSDAGSPLDMWNTWNTSNPDFTKCFONTVLVWVPCFYLWACPFYFLYLSRH 60  
 Db 1 MALRGFCSDAGSPLDMWNTWNTSNPDFTKCFONTVLVWVPCFYLWACPFYFLYLSRH 60  
 QY 61 DRGYIQTWPLNKTALGFLWLVWADLFIYSFWSRSGIFLAPVFLVSPLLGITTLLA 120

Db 61 DRGYTOMTFLNKTALGFLLLIVCWADLFYFWERSRGIPLAPVLPSPPTLLGTTLLA 120  
 QY 121 TELIOLERRKGVOSSGIMLTFWLVALVCALAILRSKIMTALKEDAQVDLFRDITFVAVPS 180  
 Db 121 TELIOLERRKGVOSSGIMLTFWLVALVCALAILRSKIMTALKEDAQVDLFRDITFVAVPS 180  
 QY 181 LLLIQLVLSFCFSDRSPFSETIHDNPNCPSESSASFLSRITFWITGLIVRGYRQPLEGSD 240  
 Db 181 LLLIQLVLSFCFSDRSPFSETIHDNPNCPSESSASFLSRITFWITGLIVRGYRQPLEGSD 240  
 QY 241 LNSLNKEDTSEQVVPVLVKNWKECAKTRKQPVVYSSKDPAPKESKVDANEVEAL 300  
 Db 241 LNSLNKEDTSEQVVPVLVKNWKECAKTRKQPVVYSSKDPAPKESKVDANEVEAL 300  
 QY 301 IVKSPQKWNPSLFVLYKTFPGPYFLMSPFFKAIHDLMMFSGPQILKLLIKFVNDTKAPD 360  
 Db 301 IVKSPQKWNPSLFVLYKTFPGPYFLMSPFFKAIHDLMMFSGPQILKLLIKFVNDTKAPD 360  
 QY 361 WQGYFYTVLLFVTFACLOTILVHOYFHCIFVSGMRKINTAVIGAYRKALVITNSARKSSTV 420  
 Db 361 WQGYFYTVLLFVTFACLOTILVHOYFHCIFVSGMRKINTAVIGAYRKALVITNSARKSSTV 420  
 QY 421 GEIVNLMSVDAORFMDLATYINMWSAPLOVILALYLLWLNIGPSVLGAVVWMLMVPYN 480  
 Db 421 GEIVNLMSVDAORFMDLATYINMWSAPLOVILALYLLWLNIGPSVLGAVVWMLMVPYN 480  
 QY 481 AVWAMTKTYQVAHMKSKDNRIKLWNEILNGIKVLKLYAWELAFKDKVLAIROEBELKVLK 540  
 Db 481 AVWAMTKTYQVAHMKSKDNRIKLWNEILNGIKVLKLYAWELAFKDKVLAIROEBELKVLK 540  
 QY 541 KSAYLSAVGTFTVCTPPLVALCTFAVYVTTIDENNILDAQTAFVSLALFNILRFPNLILP 600  
 Db 541 KSAYLSAVGTFTVCTPPLVALCTFAVYVTTIDENNILDAQTAFVSLALFNILRFPNLILP 600  
 QY 601 WYISSIVQASVLSKRLRIFLSHEELEPDSIERPVDGGTNSITVRNATFTWASDPPT 660  
 Db 601 WYISSIVQASVLSKRLRIFLSHEELEPDSIERPVDGGTNSITVRNATFTWASDPPT 660  
 QY 661 LNGITFSIPEGALVAVVGQVCGKSSLLSALLAEMDKVEGHVAIKGSVAVYVQQAWIQND 720  
 Db 661 LNGITFSIPEGALVAVVGQVCGKSSLLSALLAEMDKVEGHVAIKGSVAVYVQQAWIQND 720  
 QY 721 SIRENIFGCOLLEPYRSVIOACALLPDLLEILPSGDRTEIGEKGYNLSGGQKQVSLAR 780  
 Db 721 SIRENIFGCOLLEPYRSVIOACALLPDLLEILPSGDRTEIGEKGYNLSGGQKQVSLAR 780  
 QY 781 AVYSNADIYLFDDPLSANDAHVKGKHFENVIGPKGMLKNKTRILVTHSMSYLPQVDVIV 840  
 Db 781 AVYSNADIYLFDDPLSANDAHVKGKHFENVIGPKGMLKNKTRILVTHSMSYLPQVDVIV 840  
 QY 841 MSGGKISEMSYQELLARDGAFELRTVASTEQODAEENGVTGSGPGKEAKOMENGM 900  
 Db 841 MSGGKISEMSYQELLARDGAFELRTVASTEQODAEENGVTGSGPGKEAKOMENGM 900  
 QY 901 LVTDSAGKOLQRLSSSSSYSGDISRHNSSTAELQKAEAKKEETWKLMEADRAQTQGVKL 960  
 Db 901 LVTDSAGKOLQRLSSSSSYSGDISRHNSSTAELQKAEAKKEETWKLMEADRAQTQGVKL 960  
 QY 961 SYVDYMKAIGLFISFLSIFLPMCHNVLSALSNYMLSLWTDPIVNGTQEHTKVRLSVYG 1020  
 Db 961 SYVDYMKAIGLFISFLSIFLPMCHNVLSALSNYMLSLWTDPIVNGTQEHTKVRLSVYG 1020  
 QY 1021 ALGISOGIAVFGYSMAVSIIGGILASRCLHVDLLHLSRSPMSPFFERTPSGNLVNRFPSKEL 1080  
 Db 1021 ALGISOGIAVFGYSMAVSIIGGILASRCLHVDLLHLSRSPMSPFFERTPSGNLVNRFPSKEL 1080  
 QY 1081 DTVDSDMPEVIKMFMSGLFNIVGACIVILLATPIAAIIPPLGLIVFFVQRYVASSRQL 1140  
 Db 1081 DTVDSDMPEVIKMFMSGLFNIVGACIVILLATPIAAIIPPLGLIVFFVQRYVASSRQL 1140  
 QY 1141 KRLESVSRSPVYSHFNETLLGVSIVTRAFEEORFTHQSDLKVDENOKAYYPSIVANRWIA 1200

Db 1141 KRLESVSRSPVYSHFNETLLGVSIVTRAFEEORFTHQSDLKVDENOKAYYPSIVANRWIA 1200  
 QY 1201 VRLECVGNICVILFAALFAVISHRSLSAGLVGLSVSYSLQVTTYLNLVMSSEMETNIVA 1260  
 Db 1201 VRLECVGNICVILFAALFAVISHRSLSAGLVGLSVSYSLQVTTYLNLVMSSEMETNIVA 1260  
 QY 1261 VERLKEYSETKEAPWQIOETAPPSWPQVGRVPEPNYCLRYREDLDVFLRHINVTINGG 1320  
 Db 1261 VERLKEYSETKEAPWQIOETAPPSWPQVGRVPEPNYCLRYREDLDVFLRHINVTINGG 1320  
 QY 1321 EKVGVGTGAGKSSLTGLFRINESARGEIIGDGINIAKIGLHDLRFKTIIPQDPVLP 1380  
 Db 1321 EKVGVGTGAGKSSLTGLFRINESARGEIIGDGINIAKIGLHDLRFKTIIPQDPVLP 1380  
 QY 1381 SGSLRMNLDPPFSQYSDSEEVWTSLELAHLKDFVSALPDKLDHECAGGENLSVGQRLVCL 1440  
 Db 1381 SGSLRMNLDPPFSQYSDSEEVWTSLELAHLKDFVSALPDKLDHECAGGENLSVGQRLVCL 1440  
 QY 1441 ARALLRKTILVLDATAAVALDLETDLLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 Db 1441 ARALLRKTILVLDATAAVALDLETDLLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 QY 1501 DKGETOEYCAPSDDLQORGLFYSMAKADAGLV 1531  
 Db 1501 DKGETOEYCAPSDDLQORGLFYSMAKADAGLV 1531

RESULT 8  
 ADB20865  
 ID ADB20865 standard; protein; 1531 AA.  
 XX AC ADB20865;  
 XX DT 20-NOV-2003 (first entry)  
 XX DE MRPI based cancer related protein SEQ ID NO:678.  
 XX KW irinotecan; colorectal cancer; cervical cancer; gastric cancer;  
 KW lung cancer; ovarian cancer; pancreatic cancer; malignant glioma;  
 KW variant allele; multidrug resistance protein 1; MRPI; cytostatic.  
 XX OS Unidentified.  
 XX PN WO2003013533-A2.  
 XX PD 20-FEB-2003.  
 XX PF 23-JUL-2002; 2002WO-EP008200.  
 XX PR 23-JUL-2001; 2001EP-00117608.  
 XX PR 24-MAY-2002; 2002EP-00011710.  
 XX PA (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.  
 XX PI Heinrich G, Kerb R;  
 XX WI PI; 2003-354397/33.  
 XX PT Use of irinotecan or its derivative for preparation of a pharmaceutical  
 PT composition for treating cancer in a subject having a genome with a  
 PT variant allele comprising a multidrug resistance protein 1  
 PT polynucleotide.  
 XX PS Disclosure; SEQ ID NO 678; 100pp; English.  
 XX CC The present invention describes a method for the use of irinotecan (I) or  
 CC its derivative for the preparation of a pharmaceutical composition for  
 CC treating colorectal, cervical, gastric, lung, ovarian or pancreatic  
 CC cancer, or malignant glioma in a subject having a genome with a variant  
 CC allele which comprises a multidrug resistance protein 1 (MRPI)  
 CC polynucleotide (II). (I) has cytostatic activity. (I) or its derivative  
 CC can be used for the preparation of a pharmaceutical composition for  
 CC treating colorectal, cervical, gastric, lung, ovarian or pancreatic

CC cancer, or malignant glioma in a subject, where the subject is a human  
 CC (preferably African or Asian) or a mouse. The present sequence represents  
 CC a sequence which is used in the exemplification of the present invention.  
 XX  
 SQ Sequence 1531 AA;

Query Match 78.5%; Score 7860; DB 6; Length 1531;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MALRGFCSADGSDPLMDWNTWNTSNPDKFQNTLVVWPCFYLMWACFPFFFLYLSRH	60
DB	1	MALRGFCSADGSDPLMDWNTWNTSNPDKFQNTLVVWPCFYLMWACFPFFFLYLSRH	60
QY	61	DRGYIOWTPLNKTALGFLWLCVADLFYSPWERSRGIFLAPVLSPTLLGITLLA	120
DB	61	DRGYIOWTPLNKTALGFLWLCVADLFYSPWERSRGIFLAPVLSPTLLGITLLA	120
QY	121	TFLIQLERRKGVSOGIMLTFLVALVLCALAILRSKIMTALKEDAQVDLFRDITFVYVS	180
DB	121	TFLIQLERRKGVSOGIMLTFLVALVLCALAILRSKIMTALKEDAQVDLFRDITFVYVS	180
QY	181	LLLIQLVLSCFSDRSPLSETIHDNCPCESSASFLSRITFWITGLIVRGYRQPLEGSD	240
DB	181	LLLIQLVLSCFSDRSPLSETIHDNCPCESSASFLSRITFWITGLIVRGYRQPLEGSD	240
QY	241	LWSLNKEDTSEOVVPLVKNWKECAKTRKQPVVYSSKDPAPQKESKVDANEVEAL	300
DB	241	LWSLNKEDTSEOVVPLVKNWKECAKTRKQPVVYSSKDPAPQKESKVDANEVEAL	300
QY	301	IVKSPQKWNPSLFVLYXTGPGPYFLMSFFPKAIHDLMMFSGPQILKLIKFVNDTKAPD	360
DB	301	IVKSPQKWNPSLFVLYXTGPGPYFLMSFFPKAIHDLMMFSGPQILKLIKFVNDTKAPD	360
QY	361	WOGYFVTLVFTACLOTILVHOYPHICFVSGMRIKTAIVGAYRKALVITNSARKSTV	420
DB	361	WOGYFVTLVFTACLOTILVHOYPHICFVSGMRIKTAIVGAYRKALVITNSARKSTV	420
QY	421	GEIVNLMVSVDQRFMDLATYINNIWSAPLOVILALYLWNLGSPSLAGVAVMLMVPVN	480
DB	421	GEIVNLMVSVDQRFMDLATYINNIWSAPLOVILALYLWNLGSPSLAGVAVMLMVPVN	480
QY	481	AVWAKTKTYQVAHMKSKDNRIKLWNEILNGIKLYAWELAFKDKVLAIHQEELKVLK	540
DB	481	AVWAKTKTYQVAHMKSKDNRIKLWNEILNGIKLYAWELAFKDKVLAIHQEELKVLK	540
QY	541	KSAYLSAVGTFTWCTPFLVACTFAVVTVDENNILDAQTAFAVSLALFNILRPPLNLP	600
DB	541	KSAYLSAVGTFTWCTPFLVACTFAVVTVDENNILDAQTAFAVSLALFNILRPPLNLP	600
QY	601	MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGGINSITVRNATFTWARSDDPT	660
DB	601	MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGGINSITVRNATFTWARSDDPT	660
QY	661	LNGITFTSPEGALVAVGQVGGKSSLLSALLAEMDKVEGHVAKGSVAVYVPOQAQIWD	720
DB	661	LNGITFTSPEGALVAVGQVGGKSSLLSALLAEMDKVEGHVAKGSVAVYVPOQAQIWD	720
QY	721	SIRENIFGCOLEEPYRSVIOACALLPDLLEILPSGDRTEIGEGVNLGGQKORVSLAR	780
DB	721	SIRENIFGCOLEEPYRSVIOACALLPDLLEILPSGDRTEIGEGVNLGGQKORVSLAR	780
QY	781	AVYSNADIYLFDDPLSAVDHVGKHI FENVIGPKGMLKNKTRILVTHSMSYLPQVDVIIV	840
DB	781	AVYSNADIYLFDDPLSAVDHVGKHI FENVIGPKGMLKNKTRILVTHSMSYLPQVDVIIV	840
QY	841	MSGGKISEMGVSOELLARDGAFAPLRTYASTEODAEENGVTGVSQPGKEAKQMENG	900
DB	841	MSGGKISEMGVSOELLARDGAFAPLRTYASTEODAEENGVTGVSQPGKEAKQMENG	900
QY	901	LVTDSAGKOLQRLSSSSSYSGDISRHNSSTAEILQKAEKKEETWKLMEADKATGQVKL	960
DB	901	LVTDSAGKOLQRLSSSSSYSGDISRHNSSTAEILQKAEKKEETWKLMEADKATGQVKL	960

QY	961	SVYWDYMKAIGLFISFLSIFLFCMCHVSNVLSALSNYMLSLWTDPIVNGTOEHTKVRLSVYG	1020
DB	961	SVYWDYMKAIGLFISFLSIFLFCMCHVSNVLSALSNYMLSLWTDPIVNGTOEHTKVRLSVYG	1020
QY	1021	ALGISQGIATVFGYSMAVSIIGGILASRCLHVDLLHSILASPMSPFFERTPSGNLVNRPFSKEL	1080
DB	1021	ALGISQGIATVFGYSMAVSIIGGILASRCLHVDLLHSILASPMSPFFERTPSGNLVNRPFSKEL	1080
QY	1081	DTVDSWIMDEVIKMFWSLNFVIGACIVILLATPIAAIIIPPLGLIYFFVQRFVYVASSROL	1140
DB	1081	DTVDSWIMDEVIKMFWSLNFVIGACIVILLATPIAAIIIPPLGLIYFFVQRFVYVASSROL	1140
QY	1141	KXLESVSRSPVYSHNETLLGVSVIRAFEEQERFIHQSDLKVDENOKAYYPSIVANRWLA	1200
DB	1141	KXLESVSRSPVYSHNETLLGVSVIRAFEEQERFIHQSDLKVDENOKAYYPSIVANRWLA	1200
QY	1201	VRLECVGNCIVLFAALFAVIRHSHLSAGLVSVSYSLQVTTYLNWLVNRMSEMETNIVA	1260
DB	1201	VRLECVGNCIVLFAALFAVIRHSHLSAGLVSVSYSLQVTTYLNWLVNRMSEMETNIVA	1260
QY	1261	VERLKEYSETEKEAPWQIQETAPSSWPQVGRVFRNYCLARYREDLDFVLRHINVTINGG	1320
DB	1261	VERLKEYSETEKEAPWQIQETAPSSWPQVGRVFRNYCLARYREDLDFVLRHINVTINGG	1320
QY	1321	EKVGIVGRTGAGKSLTGLFRINESARGEIIIDGINIAKIGLHDLRPFKTIIPQDPVLF	1380
DB	1321	EKVGIVGRTGAGKSLTGLFRINESARGEIIIDGINIAKIGLHDLRPFKTIIPQDPVLF	1380
QY	1381	SGSLRMNLDPTSOYSDDEEVTSLAHLKDFVSALPDKLHCEAGGGENLSVGORQLVCL	1440
DB	1381	SGSLRMNLDPTSOYSDDEEVTSLAHLKDFVSALPDKLHCEAGGGENLSVGORQLVCL	1440
QY	1441	ARALLRKTILVLDDEATAAVALDLETDLLIQSTIRIQFEDCTVLTIAHRLNTIMDYTRVIVL	1500
DB	1441	ARALLRKTILVLDDEATAAVALDLETDLLIQSTIRIQFEDCTVLTIAHRLNTIMDYTRVIVL	1500
QY	1501	DKGEIOEGAPSDLLQORGLFYSMAKDALV	1531
DB	1501	DKGEIOEGAPSDLLQORGLFYSMAKDALV	1531

RESULT 9

ADB87954

ID ADB87954 standard; protein; 1531 AA.

AC ADB87954;

DT 04-DEC-2003 (first entry)

DE Human UGT1A1 protein sequence SEQ ID NO:678.

KW irinotecan; cancer; UGT1A1; cytostatic; topoisomerase I inhibitor;  
 KW colorectal cancer; cervical cancer; gastric cancer; lung cancer;  
 KW ovarian cancer; pancreatic cancer; malignant glioma;  
 KW uridine diphosphate glycosyltransferase1 member A1.

OS Homo sapiens.

XX WO2003013536-A2.

FN 20-FEB-2003.

PD 23-JUL-2002; 2002WO-EP008217.

PF 23-JUL-2001; 2001EP-00117608.

PR 24-MAY-2002; 2002EP-00011710.

XX (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.

XX Heinrich G, Korb R;

XX WPI; 2003-289896/28.

XX Use of irinotecan to treat cancer patient by determining if patient has  
PT variant alleles of UGT1A1 gene, administering increased/decreased amounts  
PT of irinotecan based on increased/decreased levels of UGT1A1 gene product.  
XX

PS Disclosure; SEQ ID NO 678; 107pp; English.

XX The invention relates to the novel use of irinotecan to treat a patient  
CC suffering from cancer. This involves determining if the patient has one  
CC or more variant alleles of the UGT1A1 gene, and if the patient has one or  
CC more of such variant alleles, irinotecan is administered in an increased  
CC or decreased amount in comparison to the amount that is administered  
CC without regard to the patient's alleles in the UGT1A1 gene. The invention  
CC has cytostatic activity. A composition of the invention acts as a  
CC topoisomerase I inhibitor. The method is useful for treating a patient,  
CC an animal e.g. mouse or a human, preferably African or Asian, suffering  
CC from cancer such as colorectal, cervical, gastric cancer, lung, ovarian,  
CC pancreatic cancer or malignant glioma. The present sequence is udes in  
CC the exemplification of the invention.  
XX

SQ Sequence 1531 AA;

Query Match 78.5%; Score 7860; DB 7; Length 1531;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MALRGFCSADGSDPLMDWNVWNTSNPDKFQNTLVVWPCFVYLWACFPFVLYLSRH	60
DB	1	MALRGFCSADGSDPLMDWNVWNTSNPDKFQNTLVVWPCFVYLWACFPFVLYLSRH	60
QY	61	DRGYQMTPLNKTALGELLIVCWADLFYFWSRSGIFLAPVFLVSPDLLGTTLLA	120
DB	61	DRGYQMTPLNKTALGELLIVCWADLFYFWSRSGIFLAPVFLVSPDLLGTTLLA	120
QY	121	TFLIQLERKGVQSSGIMLTFWLVALVCAILRLSKIMTALKEDAQVDLFRDITVYVFS	180
DB	121	TFLIQLERKGVQSSGIMLTFWLVALVCAILRLSKIMTALKEDAQVDLFRDITVYVFS	180
QY	181	LLLIQLVLSFCSDRSPFLSETIHDNPPCPSSASFLSRITFWIITGLIVRGVROPLEGSD	240
DB	181	LLLIQLVLSFCSDRSPFLSETIHDNPPCPSSASFLSRITFWIITGLIVRGVROPLEGSD	240
QY	241	LWSLNKEDTSEQVVPVLVKNMKECAKTRKQPVVYSSKDPAPKESKVDANEVEAL	300
DB	241	LWSLNKEDTSEQVVPVLVKNMKECAKTRKQPVVYSSKDPAPKESKVDANEVEAL	300
QY	301	IVKSPQKWNPSLFKVLKTFGYPFLMSFPFKAIDHLMFSGPQILKLLIKFVNDTKAPD	360
DB	301	IVKSPQKWNPSLFKVLKTFGYPFLMSFPFKAIDHLMFSGPQILKLLIKFVNDTKAPD	360
QY	361	WQGYFVTVLLFVTACLOTLVHLYPHI CFVSGMRKTAIVGAVYRKALVITNSARKSTV	420
DB	361	WQGYFVTVLLFVTACLOTLVHLYPHI CFVSGMRKTAIVGAVYRKALVITNSARKSTV	420
QY	421	GEIVNLMSVDAQRFMDLATYINNIWSAPLQVILALYLNLNGLSPSLAGVAVVWLMPVN	480
DB	421	GEIVNLMSVDAQRFMDLATYINNIWSAPLQVILALYLNLNGLSPSLAGVAVVWLMPVN	480
QY	481	AVAMKTKTYQVAHMKSKDNRIKLMEINLNGIKVLKYAWELAFKDKVLAIRQEBELKVLK	540
DB	481	AVAMKTKTYQVAHMKSKDNRIKLMEINLNGIKVLKYAWELAFKDKVLAIRQEBELKVLK	540
QY	541	KSAYLSAVCTFTWCTPFLVALCTFAVYVITDENNILDQAOTAPVSLAFNLIIRPPLNITLP	600
DB	541	KSAYLSAVCTFTWCTPFLVALCTFAVYVITDENNILDQAOTAPVSLAFNLIIRPPLNITLP	600
QY	601	WVISSIVQASVSLKRLIRIFLSHEELEPDSIERPPVKDGGGTSITVRNATFTWASDDPT	660
DB	601	WVISSIVQASVSLKRLIRIFLSHEELEPDSIERPPVKDGGGTSITVRNATFTWASDDPT	660
QY	661	LNGITFISIEGALVAVVGQVCGCKSLLSALLAEMDKVEGHVAIKGSVAVVPPQAWIQND	720
DB	661	LNGITFISIEGALVAVVGQVCGCKSLLSALLAEMDKVEGHVAIKGSVAVVPPQAWIQND	720

QY	721	SIRENIFGCOLLEPYRSVSIQACALLPDLLEILPSGDRTEIGEKGVLNLSGGQKQVSLAR	780
DB	721	SIRENIFGCOLLEPYRSVSIQACALLPDLLEILPSGDRTEIGEKGVLNLSGGQKQVSLAR	780
QY	781	AVYSNADITYLFDPLSAVDHVGKHI FENVIGPKGMLKNKTRILVTHSMYSYLPQVDVIIV	840
DB	781	AVYSNADITYLFDPLSAVDHVGKHI FENVIGPKGMLKNKTRILVTHSMYSYLPQVDVIIV	840
QY	841	MSGGKISMGVSQELLARDGAFABFLRYASTEQDQDAEENGVTGSPGPKAKOMENGM	900
DB	841	MSGGKISMGVSQELLARDGAFABFLRYASTEQDQDAEENGVTGSPGPKAKOMENGM	900
QY	901	LVTDSAGLQORQLSSSSSSSGDISRHNHSTAELQAKAEKBEETWKLMEADKAQTCQVXL	960
DB	901	LVTDSAGLQORQLSSSSSSSGDISRHNHSTAELQAKAEKBEETWKLMEADKAQTCQVXL	960
QY	961	SVYDMYKAIGLFTSIFLIFPMCNHVSALASNYWLSLWTDPIVNGTQEHTKVRLSVYG	1020
DB	961	SVYDMYKAIGLFTSIFLIFPMCNHVSALASNYWLSLWTDPIVNGTQEHTKVRLSVYG	1020
QY	1021	ALGISQGIATVFGYSMAVISGIGILASRCILHVDLLHSILRSPMSFFERTPSGNLVNRFPSKEL	1080
DB	1021	ALGISQGIATVFGYSMAVISGIGILASRCILHVDLLHSILRSPMSFFERTPSGNLVNRFPSKEL	1080
QY	1081	DTVDSMIPEVIMKFMGSLFNVIACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSROL	1140
DB	1081	DTVDSMIPEVIMKFMGSLFNVIACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSROL	1140
QY	1141	KRLESVSRSPVYSHNETLLGVSVIRAFEEQERFIHQSDLKVDENOKAYYPSIVANRWLA	1200
DB	1141	KRLESVSRSPVYSHNETLLGVSVIRAFEEQERFIHQSDLKVDENOKAYYPSIVANRWLA	1200
QY	1201	VRLECVGNCIVLFAALFAVIRHSLASGLSVSLSQVTTLYLNLVMSSEMETNIVA	1260
DB	1201	VRLECVGNCIVLFAALFAVIRHSLASGLSVSLSQVTTLYLNLVMSSEMETNIVA	1260
QY	1261	VERLKEYSETEKEAPWQIQTAPPSSWPQVGRVEFRNYCLRYREDLDFVLRHINVTINGG	1320
DB	1261	VERLKEYSETEKEAPWQIQTAPPSSWPQVGRVEFRNYCLRYREDLDFVLRHINVTINGG	1320
QY	1321	EKVGIVGRTGAGKSLTGLFRINESARGEIIDGINIAKIGLHDLRFKTIIPQDPVLF	1380
DB	1321	EKVGIVGRTGAGKSLTGLFRINESARGEIIDGINIAKIGLHDLRFKTIIPQDPVLF	1380
QY	1381	SGSLRMNLDPSOYSDDEEVTSLAHLKDFVSALPDKLDHECAGGENISVGQRLVCL	1440
DB	1381	SGSLRMNLDPSOYSDDEEVTSLAHLKDFVSALPDKLDHECAGGENISVGQRLVCL	1440
QY	1441	ARALLRKTILVLDATAAVALDLETDLLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL	1500
DB	1441	ARALLRKTILVLDATAAVALDLETDLLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL	1500
QY	1501	DKGETQYGAPEGLLQORGLFYSMAKADGLV	1531
DB	1501	DKGETQYGAPEGLLQORGLFYSMAKADGLV	1531

RESULT 10  
ADB96937

ID ADB96937 standard; protein; 1531 AA.

XX AC ADB96937;

XX DT 04-DEC-2003 (first entry)

XX DE Human MDR1 related protein sequence SEQ ID NO:678.

XX KW irinotecan; colorectal cancer; cervical cancer; gastric cancer;

XX KW lung cancer; ovarian cancer; pancreatic cancer; malignant glioma;

XX KW multidrug resistance 1; MDR1; cytostatic; human; CYP3A5; MRP1; MDR1;

XX TOPI.

OS	Homo sapiens.	QY	481	AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKLYAWELAFKDKVLAIRQBELKVLK	540
XX					
PN	W02003013537-A2.	Db	481	AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKLYAWELAFKDKVLAIRQBELKVLK	540
XX					
PD	20-FEB-2003.	QY	541	KSAYLSAVGTFTWCTPFLVALCTFAVVTIDENNILDAQTAFAVSLALFNILRPLNLP	600
XX					
PF	23-JUL-2002; 2002WO-EP008218.	Db	541	KSAYLSAVGTFTWCTPFLVALCTFAVVTIDENNILDAQTAFAVSLALFNILRPLNLP	600
XX					
PR	23-JUL-2001; 2001EP-00117608.	QY	601	MVISIVQASVSLKRLRIFLSHEBLEPDSIERRPVKDGSGTNSITVRNATFTWARSDEPT	660
PR	24-MAY-2002; 2002EP-00011710.	Db	601	MVISIVQASVSLKRLRIFLSHEBLEPDSIERRPVKDGSGTNSITVRNATFTWARSDEPT	660
XX					
PA	(EPID-) EPIDAURUS BIOTECHNOLOGIE AG.	QY	661	LNGITFISPEGALVAVVGVCGKSSLLSALLAEWDMVEGHVAIKGSVAVVPOQAWTQND	720
XX					
PI	Heinrich G, Kerb R;	Db	661	LNGITFISPEGALVAVVGVCGKSSLLSALLAEWDMVEGHVAIKGSVAVVPOQAWTQND	720
XX					
DR	WPI; 2003-268145/26.	QY	721	SLRENILFGCQLEBPYRSVIQACALLPDLEILPSGDRTEIGEGVNLSSGQKORVSLAR	780
XX					
XX	New use of irinotecan for preparation of pharmaceutical compositions for	Db	721	SLRENILFGCQLEBPYRSVIQACALLPDLEILPSGDRTEIGEGVNLSSGQKORVSLAR	780
PT	treating cancer in subject having genome with variant allele comprising				
PT	multidrug resistance 1 polynucleotide.				
XX					
PS	Disclosure; SEQ ID NO 678; 130pp; English.	QY	781	AVYSNADIYLFDDPLSAVDHVGKHFENVIKPGKMLKNKTRILVTHSMYSYLPQVDVIIV	840
XX					
XX	The invention relates to the novel use of irinotecan or its derivative	Db	781	AVYSNADIYLFDDPLSAVDHVGKHFENVIKPGKMLKNKTRILVTHSMYSYLPQVDVIIV	840
CC	for the preparation of pharmaceutical compositions for treating				
CC	colorectal, cervical, gastric, lung, ovarian or pancreatic cancer, or	QY	841	MSGGKISEMSSYQELLARDGAFARFLRTYASTEOQDAENGVTGVSQPKGKAKOMENGM	900
CC	malignant glioma in a subject having a genome with a variant allele which	Db	841	MSGGKISEMSSYQELLARDGAFARFLRTYASTEOQDAENGVTGVSQPKGKAKOMENGM	900
CC	comprises a multidrug resistance 1 (MDR1) polynucleotide. A composition				
CC	of the invention has cytostatic activity. The invention is useful for the	QY	901	LVTDSACKQLQRLQSSSSSYSGDTSRHNHSTAEIQAQAEKEETWKLMEADKAGTQGVKL	960
CC	preparation of pharmaceutical compositions for treating colorectal,	Db	901	LVTDSACKQLQRLQSSSSSYSGDTSRHNHSTAEIQAQAEKEETWKLMEADKAGTQGVKL	960
CC	cervical, gastric, lung, ovarian or pancreatic cancer, or malignant				
CC	glioma in a subject (preferably human, more preferably African or Asian)	QY	961	SVYDYMKAIGLFIISFLSIFLFCNHNVSALASNYLSLWTDPIVNGTQSHTKVRLSVYG	1020
CC	or a mouse. The present sequence is used in the exemplification of the	Db	961	SVYDYMKAIGLFIISFLSIFLFCNHNVSALASNYLSLWTDPIVNGTQSHTKVRLSVYG	1020
XX					
SQ	Sequence 1531 AA;	QY	1021	ALGISQGIAGVFGSMAYSIGGILASRCLHVDLHLSILRSPMSFFERTPSGNLVNRPSEL	1080
	Query Match 78.5%; Score 7860; DB 7; Length 1531;	Db	1021	ALGISQGIAGVFGSMAYSIGGILASRCLHVDLHLSILRSPMSFFERTPSGNLVNRPSEL	1080
	Best Local Similarity 100.0%; Pred. No. 0;				
	Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	QY	1081	DTVDSMTPEVVKPMGSLFNVIACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQL	1140
		Db	1081	DTVDSMTPEVVKPMGSLFNVIACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQL	1140
QY	1 MAIRGFCSDGSDPLDNDVNTWNTSNPDFTKCFQNTVLVWVPCFVLMWACFPFFFLYLSRH	QY	1141	KRLSVSRSPVYSHENETLLGVSVIRAFEEQERFIHQSDLKVDENQKAYYPSIVANRWLA	1200
Db	1 MAIRGFCSDGSDPLDNDVNTWNTSNPDFTKCFQNTVLVWVPCFVLMWACFPFFFLYLSRH	Db	1141	KRLSVSRSPVYSHENETLLGVSVIRAFEEQERFIHQSDLKVDENQKAYYPSIVANRWLA	1200
QY	61 DRGYIQTPLNKTALGFLMWVCWADLFYSFWSRSRGIFLAPVFLVSPFLGTTLLA	QY	1201	VRLECVGNCIVLFAALFAVIRSHLSAGLVGLSVSYSLQVTTVYLNMLVRMSSEMETNIVA	1260
Db	61 DRGYIQTPLNKTALGFLMWVCWADLFYSFWSRSRGIFLAPVFLVSPFLGTTLLA	Db	1201	VRLECVGNCIVLFAALFAVIRSHLSAGLVGLSVSYSLQVTTVYLNMLVRMSSEMETNIVA	1260
QY	121 TFLIQLERRKGVQSSGIMLTFWLVALVCAALILRSKIMTALKEDAQVDLFRDITFYVYFS	QY	1261	VERLKEYSETEKAPWQIOETAPPSPWPQVGRVFRNRYCLRYREDLDFVLRHINVTINGG	1320
Db	121 TFLIQLERRKGVQSSGIMLTFWLVALVCAALILRSKIMTALKEDAQVDLFRDITFYVYFS	Db	1261	VERLKEYSETEKAPWQIOETAPPSPWPQVGRVFRNRYCLRYREDLDFVLRHINVTINGG	1320
QY	181 LLLIQLVLSCSFSDRPLSETHDNPCESSASFLSRITFWITGLIVRGYRQPLEGSD	QY	1321	EKVGIVGRTGAGKSSLTGLFRINESAGEEIIIDGINIAKIGLHDLRFKTIIPQDPVLF	1380
Db	181 LLLIQLVLSCSFSDRPLSETHDNPCESSASFLSRITFWITGLIVRGYRQPLEGSD	Db	1321	EKVGIVGRTGAGKSSLTGLFRINESAGEEIIIDGINIAKIGLHDLRFKTIIPQDPVLF	1380
QY	241 LWSLNKEDTSEQVPVLVKNWKKCAKTRKQPVKVYSSKDPAPKESKVDANEVEAL	QY	1381	SGSLRMNLDPFQSYSDSEEVWTSLELAHLKDFVSALPDKLDHECAEGENLSVGQRLVCL	1440
Db	241 LWSLNKEDTSEQVPVLVKNWKKCAKTRKQPVKVYSSKDPAPKESKVDANEVEAL	Db	1381	SGSLRMNLDPFQSYSDSEEVWTSLELAHLKDFVSALPDKLDHECAEGENLSVGQRLVCL	1440
QY	301 IVKSPQKEWNSLFLKVLVKTGPFVLMFFPKAIHDLMMFSGPQILKLLIKFVNDTKAPD	QY	1441	ARALLRKTLLVLDEATAVDLETDLLIOSTIRTOQEDCTVLTIAHRLNTIMDYTRVIL	1500
Db	301 IVKSPQKEWNSLFLKVLVKTGPFVLMFFPKAIHDLMMFSGPQILKLLIKFVNDTKAPD	Db	1441	ARALLRKTLLVLDEATAVDLETDLLIOSTIRTOQEDCTVLTIAHRLNTIMDYTRVIL	1500
QY	361 WQGYEYTVLLFVTAQLVHLQVHFICFVSGMRKTAIVGAVYRKALVITNSARKSSTV	QY	1501	DKGEIYEGAPSDLLQORGLFYSMKADAGLV	1531
Db	361 WQGYEYTVLLFVTAQLVHLQVHFICFVSGMRKTAIVGAVYRKALVITNSARKSSTV	Db	1501	DKGEIYEGAPSDLLQORGLFYSMKADAGLV	1531
QY	421 GEIVNLSVDAQRFMDLATYINMIWSAPLQVILALYLMLNGLSPVLGAVVMVLPVN	QY			
Db	421 GEIVNLSVDAQRFMDLATYINMIWSAPLQVILALYLMLNGLSPVLGAVVMVLPVN	Db			

RESULT 11  
AD892128  
ID ADB92128 standard; protein; 1531 AA.  
XX AC ADB92128;  
XX DT 04-DEC-2003 (first entry)  
XX DE Human MDR1 related protein sequence SEQ ID NO:678.  
XX KW irinotecan; colorectal cancer; cervical cancer; gastric cancer;  
XX KW lung cancer; ovarian cancer; pancreatic cancer; malignant glioma;  
XX KW multidrug resistance 1; MDR1; cytostatic; human; UGT1A1; MRP1; TOP1.  
XX OS Homo sapiens.  
XX PN WO2003013535-A2.  
XX PD 20-FEB-2003.  
XX PF 23-JUL-2002; 2002WO-EP008220.  
XX PR 23-JUL-2001; 2001EP-00117608.  
XX PR 24-MAY-2002; 2002EP-00011710.  
XX PA (EPID-) EPIDAURUS BIOTECHNOLOGIE AG.  
XX PI Heinrich G, Kerb R;  
XX WPI; 2003-342400/32.  
XX  
XX New use of irinotecan for preparation of pharmaceutical compositions for  
XX treating cancer in subject having genome with variant allele comprising  
XX multidrug resistance 1 polynucleotide.  
XX Disclosure; SEQ ID NO 678; 104pp; English.  
XX  
XX The invention relates to a novel use of irinotecan or its derivative for  
XX the preparation of a pharmaceutical composition for treating colorectal,  
XX cervical, gastric, lung, ovarian or pancreatic cancer, or malignant  
XX glioma in a subject having a genome with a variant allele which comprises  
XX a multidrug resistance 1 (MDR1) polynucleotide. A composition of the  
XX invention has cytostatic activity. The present sequence is used in the  
XX exemplification of the invention.  
XX  
XX Sequence 1531 AA;  
SQ  
Query Match 78.5%; Score 7860; DB 7; Length 1531;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MALRGFCSADGSDPLMDNMTNTNSPDFTKCFQNTVLVWPCFYLWACFPFYLGRH 60  
DB 1 MALRGFCSADGSDPLMDNMTNTNSPDFTKCFQNTVLVWPCFYLWACFPFYLGRH 60  
QY 61 DRGYQMTPLNKTALGFLLMVWADLFYFWSRSGIFLAPFLVSPDLLGTTLLA 120  
DB 61 DRGYQMTPLNKTALGFLLMVWADLFYFWSRSGIFLAPFLVSPDLLGTTLLA 120  
QY 121 TFLIQLERRKGVSQSGIMLTFWLVVALVCAALILRSKIMTALKEDAQVDLFRDITYVYFS 180  
DB 121 TFLIQLERRKGVSQSGIMLTFWLVVALVCAALILRSKIMTALKEDAQVDLFRDITYVYFS 180  
QY 181 LLLIQLVLSGSDRSPLETHDNPCESSASFLSRITFWITGLIVRGVROPLEGSD 240  
DB 181 LLLIQLVLSGSDRSPLETHDNPCESSASFLSRITFWITGLIVRGVROPLEGSD 240  
QY 241 LMSLNKEDTSEQVPLVKNWKECAKTRKQPKVYSSKDPAPKSSKVDANEVEAL 300  
DB 241 LMSLNKEDTSEQVPLVKNWKECAKTRKQPKVYSSKDPAPKSSKVDANEVEAL 300  
QY 301 IVKSPQKWNPSLFKVLKTFPGYFLMSFFFAKIHDLMMFSGPQILKLLIKFVNDTKAPD 360

DB 301 IVKSPQKWNPSLFKVLKTFPGYFLMSFFFAKIHDLMMFSGPQILKLLIKFVNDTKAPD 360  
QY 361 WQGYFTVLLFVTACLOTLVLHOYPHICFVSGMRKTAIVIGAVYKALVITNSARKSSTV 420  
DB 361 WQGYFTVLLFVTACLOTLVLHOYPHICFVSGMRKTAIVIGAVYKALVITNSARKSSTV 420  
QY 421 GEIVNLSVDAQRFMDLATYINMIWSAPLQVILALYLLWNLGSPVLAGVAVMVLMPVN 480  
DB 421 GEIVNLSVDAQRFMDLATYINMIWSAPLQVILALYLLWNLGSPVLAGVAVMVLMPVN 480  
QY 481 AVMAKTKTYQVAHMKSKDNRIKLNNELINGIKVLYAWELAFKDKVLAIRQEBELKVLK 540  
DB 481 AVMAKTKTYQVAHMKSKDNRIKLNNELINGIKVLYAWELAFKDKVLAIRQEBELKVLK 540  
QY 541 KSAVLSAVGTFTWCTPFLVALCTFAVVVTDENNILDAQTAFAVSLAFNIRFPLNILP 600  
DB 541 KSAVLSAVGTFTWCTPFLVALCTFAVVVTDENNILDAQTAFAVSLAFNIRFPLNILP 600  
QY 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWASDDPT 660  
DB 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWASDDPT 660  
QY 661 LNGITFSIPEGALVAVGVGCGKSSLSALLAEMDKVEGHVAIKGSVAYVPOQAWIQND 720  
DB 661 LNGITFSIPEGALVAVGVGCGKSSLSALLAEMDKVEGHVAIKGSVAYVPOQAWIQND 720  
QY 721 SLRENILFGCOLLEPYRSVIOACALLPOLEILPSGDRTEIGEKGVLNLSGGQKQVSLAR 780  
DB 721 SLRENILFGCOLLEPYRSVIOACALLPOLEILPSGDRTEIGEKGVLNLSGGQKQVSLAR 780  
QY 781 AVYSNADTYLDDPLSADVAHVGHIFENVIGPKMLKNKTRILVTHSNLYLPQVDVIV 840  
DB 781 AVYSNADTYLDDPLSADVAHVGHIFENVIGPKMLKNKTRILVTHSNLYLPQVDVIV 840  
QY 841 MSGGKISEMSYQELLARDGAFAEFLRTYASTEQDABENGVTGVSFGPKAKOMENGM 900  
DB 841 MSGGKISEMSYQELLARDGAFAEFLRTYASTEQDABENGVTGVSFGPKAKOMENGM 900  
QY 901 LVTDSAGKOLQKQSSSSSSSYSGDISRHNSHTAELQKAEKKEETWKLMEADKAQGOVKL 960  
DB 901 LVTDSAGKOLQKQSSSSSSSYSGDISRHNSHTAELQKAEKKEETWKLMEADKAQGOVKL 960  
QY 961 SVYDYMKAIGLIFISFLSIFLFCMCHVSNALSNYLSLWTDPIVNGTQEHKTVRLSVYG 1020  
DB 961 SVYDYMKAIGLIFISFLSIFLFCMCHVSNALSNYLSLWTDPIVNGTQEHKTVRLSVYG 1020  
QY 1021 ALGISQGIATVFGYSNAVSIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFSEL 1080  
DB 1021 ALGISQGIATVFGYSNAVSIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFSEL 1080  
QY 1081 DTVDSMIPEVIMKFMGSLFNVIGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQL 1140  
DB 1081 DTVDSMIPEVIMKFMGSLFNVIGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQL 1140  
QY 1141 KRLESVRSRSPVSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENOKAYPSIVANRWLA 1200  
DB 1141 KRLESVRSRSPVSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENOKAYPSIVANRWLA 1200  
QY 1201 VRLECVCNCIVLFAALFAVISRHSLSAGVLSVSYSLQVTTYLANLVRMSSEMETNIVA 1260  
DB 1201 VRLECVCNCIVLFAALFAVISRHSLSAGVLSVSYSLQVTTYLANLVRMSSEMETNIVA 1260  
QY 1261 VERLKEYSETEKEAPWQIQETAPSSWPQGVGFENYCLRYREDLDFVLRHINTVINGG 1320  
DB 1261 VERLKEYSETEKEAPWQIQETAPSSWPQGVGFENYCLRYREDLDFVLRHINTVINGG 1320  
QY 1321 EKVGIVGRTGAGKSSLTGLFRINESABEIIIDGINIAKIGLHDLRFKTIIPDPVLF 1380  
DB 1321 EKVGIVGRTGAGKSSLTGLFRINESABEIIIDGINIAKIGLHDLRFKTIIPDPVLF 1380  
QY 1381 SGLSRMNLDPFSQVSDDEEVTSLAHLKDFVSALPDKLDHECAGGENLSVQRLVCL 1440  
DB 1381 SGLSRMNLDPFSQVSDDEEVTSLAHLKDFVSALPDKLDHECAGGENLSVQRLVCL 1440

QY 1441 ARALLRKTILVLDEATAAVDLEDDLIQSTIOTPEBDCVLTIAHRLNTIMDYTRVIVL 1500  
 DB 1441 ARALLRKTILVLDEATAAVDLEDDLIQSTIOTPEBDCVLTIAHRLNTIMDYTRVIVL 1500  
 QY 1501 DKGEIQEYGAPSDLLQORGLFYSMAXDAGLV 1531  
 DB 1501 DKGEIQEYGAPSDLLQORGLFYSMAXDAGLV 1531

RESULT 12  
 ADD44764  
 ID ADD44764 standard; protein; 1531 AA.  
 XX  
 AC ADD44764;  
 XX  
 DT 29-JAN-2004 (first entry)  
 XX  
 DE Human Protein P33527, SEQ ID NO 10193.  
 XX  
 KW Human; pain; neuronal tissue; gene therapy;  
 KW spinal segmental nerve injury; chronic constriction injury; CCI;  
 KW spared nerve injury; SNI; Chung.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2003016475-A2.  
 XX  
 PD 27-FEB-2003.  
 XX  
 PF 14-AUG-2002; 2002WO-US025765.  
 XX  
 PR 14-AUG-2001; 2001US-0312147P.  
 PR 01-NOV-2001; 2001US-0346382P.  
 PR 26-NOV-2001; 2001US-0333347P.  
 XX  
 XX (GEO) GEN HOSPITAL CORP.  
 PA (FARB) BAYER AG.  
 XX  
 XX  
 PI Woolf C, D'urso D, Befort K, Costigan M;  
 XX  
 XX WPI; 2003-268312/26.  
 DR GENBANK; P33527.  
 XX  
 XX New composition comprising two or more isolated polypeptides, useful for  
 PT preparing a medicament for treating pain in an animal.  
 PS  
 PS Claim 1; Page; 1017pp; English.

CC The invention discloses a composition comprising two or more isolated rat  
 CC or human polynucleotides or a polynucleotide which represents a fragment,  
 CC derivative or allelic variation of the nucleic acid sequence. Also  
 CC claimed are a vector comprising the novel polynucleotide, a host cell  
 CC comprising the vector, a method for identifying a nucleotide sequence  
 CC which is differentially regulated in an animal subjected to pain and a  
 CC kit to perform the method, an array, a method for identifying an agent  
 CC that increases or decreases the expression of the polynucleotide sequence  
 CC that is differentially expressed in neuronal tissue of a first animal  
 CC subjected to pain, a method for identifying a compound which regulates  
 CC the expression of a polynucleotide sequence which is differentially  
 CC expressed in an animal subjected to pain, a method for identifying a  
 CC compound that regulates the activity of one or more of the  
 CC polynucleotides, a method for producing a pharmaceutical composition, a  
 CC method for identifying a compound or small molecule that regulates the  
 CC activity in an animal of one or more of the polypeptides given in the  
 CC specification, a method for identifying a compound useful in treating  
 CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
 CC therapy). The sequence presented is a human protein (shown in Table 2 of  
 CC the specification) which is differentially expressed during pain. Note:

CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 1531 AA;  
 Query Match 78.5%; Score 7860; DB 7; Length 1531;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MALRGFCSDGSDPLMDNVTWNTSNPDFTKCFQNTLVWVPCFVLMACFPFYLYLSRH 60  
 DB 1 MALRGFCSDGSDPLMDNVTWNTSNPDFTKCFQNTLVWVPCFVLMACFPFYLYLSRH 60  
 QY 61 DRGYIQMTPLNKTALGFLMWIVCWADLFYSFWSERSGIFLAPVFLVSPITLLGTTLLA 120  
 DB 61 DRGYIQMTPLNKTALGFLMWIVCWADLFYSFWSERSGIFLAPVFLVSPITLLGTTLLA 120  
 QY 121 TFLIQLEERRKGVOSSGIMLTFWLVALVCALAILRSKIMTALKEDAQVDFRDIITYVYVS 180  
 DB 121 TFLIQLEERRKGVOSSGIMLTFWLVALVCALAILRSKIMTALKEDAQVDFRDIITYVYVS 180  
 QY 181 LLLIQVLSCFSDRSPLEFSETHDNPCESSASFLSRITFWITGLIVRGVROPLEGSD 240  
 DB 181 LLLIQVLSCFSDRSPLEFSETHDNPCESSASFLSRITFWITGLIVRGVROPLEGSD 240  
 QY 241 LWSLNKEDTSQVVPVVLVKNWKECAKTRKQPVKYVYSSKDPAPQKSSKVDANEEVEAL 300  
 DB 241 LWSLNKEDTSQVVPVVLVKNWKECAKTRKQPVKYVYSSKDPAPQKSSKVDANEEVEAL 300  
 QY 301 IVKSPQKEWNPFLFKVLYKTFPGYFLMSFFPKAHDLMFSGPQILKLLIKFVNDTKAPD 360  
 DB 301 IVKSPQKEWNPFLFKVLYKTFPGYFLMSFFPKAHDLMFSGPQILKLLIKFVNDTKAPD 360  
 QY 361 WQGYFYTVLLFVTACLOTLVLHQYFHICFVSGMERIKTAVIGAVYRKALVITNSARKSTV 420  
 DB 361 WQGYFYTVLLFVTACLOTLVLHQYFHICFVSGMERIKTAVIGAVYRKALVITNSARKSTV 420  
 QY 421 GEIVNLMSVDAQRFMDLATYINMIWSAPLQVILALYLLNMLGSPVLGAVVWVMPVN 480  
 DB 421 GEIVNLMSVDAQRFMDLATYINMIWSAPLQVILALYLLNMLGSPVLGAVVWVMPVN 480  
 QY 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLYAWELAFKDKVLAIRQELKVLK 540  
 DB 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLYAWELAFKDKVLAIRQELKVLK 540  
 QY 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVITDENNILDQAOTAFVSLALFNILRFPILNP 600  
 DB 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVITDENNILDQAOTAFVSLALFNILRFPILNP 600  
 QY 601 MVISSIVQASVSLKRLRIFLUSHELEPDSIERRPVKDGSGTNSITVRNATFTWARSDDPT 660  
 DB 601 MVISSIVQASVSLKRLRIFLUSHELEPDSIERRPVKDGSGTNSITVRNATFTWARSDDPT 660  
 QY 661 LINGITFIPGALVAVVGVQVCGCKSSLLSALLAEMDKVEGHVAIKGSVAVVPOQAWTQND 720  
 DB 661 LINGITFIPGALVAVVGVQVCGCKSSLLSALLAEMDKVEGHVAIKGSVAVVPOQAWTQND 720  
 QY 721 SLRENILFGCLBEPYRVSIVIOCALLPDLIELIPSGDRTEIGEGVNLSSGGQKQVSLAR 780  
 DB 721 SLRENILFGCLBEPYRVSIVIOCALLPDLIELIPSGDRTEIGEGVNLSSGGQKQVSLAR 780  
 QY 781 AVYSNADIYLFDDPLSAVDARVGHKIFENVIGPKMLKNKTRILVTHSMYSYLPQVDVII 840  
 DB 781 AVYSNADIYLFDDPLSAVDARVGHKIFENVIGPKMLKNKTRILVTHSMYSYLPQVDVII 840  
 QY 841 MSGGKISEMSYQELLARDGAFAEFLRTYASTEQEDAENGVTGSGPGKEAKQOMENGM 900  
 DB 841 MSGGKISEMSYQELLARDGAFAEFLRTYASTEQEDAENGVTGSGPGKEAKQOMENGM 900  
 QY 901 LVTDTSAGKQLQROLSSSSSYSGDISRHHNSTAELOKAEKKEETWKLMEADKQOTGVKL 960  
 DB 901 LVTDTSAGKQLQROLSSSSSYSGDISRHHNSTAELOKAEKKEETWKLMEADKQOTGVKL 960

QY 961 SYVDYMKAGLIFISFLIFPMCNHVSALASNYWLSLWTDPIVNGTQEHKVLRSVYG 1020  
 DB 961 SYVDYMKAGLIFISFLIFPMCNHVSALASNYWLSLWTDPIVNGTQEHKVLRSVYG 1020  
 QY 1021 ALGISQGIATVFGYSMAVSIIGGILASRCLHVDLHLSILSPMSFFERTPSGNLVNRFSEL 1080  
 DB 1021 ALGISQGIATVFGYSMAVSIIGGILASRCLHVDLHLSILSPMSFFERTPSGNLVNRFSEL 1080  
 QY 1081 DTVDMSIPEVIMKFMGSLFNVIACIVILLATPIAAIIIPPLGLIYFFVQRPYVASSRQL 1140  
 DB 1081 DTVDMSIPEVIMKFMGSLFNVIACIVILLATPIAAIIIPPLGLIYFFVQRPYVASSRQL 1140  
 QY 1141 KELESVSRSPVSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENQKAYPSIVANRWLA 1200  
 DB 1141 KELESVSRSPVSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENQKAYPSIVANRWLA 1200  
 QY 1201 VRLECVGNCIVLFAALFAVISHSLISAGLVGLSVSYSLQVTTYLNWLVMSSEMETNIVA 1260  
 DB 1201 VRLECVGNCIVLFAALFAVISHSLISAGLVGLSVSYSLQVTTYLNWLVMSSEMETNIVA 1260  
 QY 1261 VERLKEYSETEKAPWQIQTAPPSSWPOVGRVFNRYCLRYREDLDFVLRHINVTINGG 1320  
 DB 1261 VERLKEYSETEKAPWQIQTAPPSSWPOVGRVFNRYCLRYREDLDFVLRHINVTINGG 1320  
 QY 1321 EKVIGVGRGACKSSITLGLFRINESAEGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
 DB 1321 EKVIGVGRGACKSSITLGLFRINESAEGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
 QY 1381 SGLSRMNLDPFQSYSDVEWTSLELAHLKDFVSALPKLDHECARGGENLSVGQRQLVCL 1440  
 DB 1381 SGLSRMNLDPFQSYSDVEWTSLELAHLKDFVSALPKLDHECARGGENLSVGQRQLVCL 1440  
 QY 1441 APALARUKTKILVDBATAVDLETDLLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 DB 1441 APALARUKTKILVDBATAVDLETDLLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 QY 1501 DKGEIOEYGAPSDLLQORGLFYSMADAGLV 1531  
 DB 1501 DKGEIOEYGAPSDLLQORGLFYSMADAGLV 1531  
 RESULT 13  
 ADN97111  
 ID ADN97111 standard; protein; 1531 AA.  
 AC ADN97111;  
 XX ADN97111;  
 DT 01-JUL-2004 (first entry)  
 DE MRP1 protein.  
 KW multiple drug resistance protein; MRP; Drosophila melanogaster;  
 KW Anopheles gambiae; insecticide.  
 XX Homo sapiens.  
 XX WO2004029088-A2.  
 XX 08-APR-2004.  
 XX 25-SEP-2003; 2003WO-EP012400.  
 XX 26-SEP-2002; 2002US-0413469P.  
 XX (INSP ) INST PASTEUR.  
 XX (CNRS ) CENT NAT RECH SCI.  
 XX PI Roth CW, Brey PT, Holm I, Graillies M, Rzhetsky A;  
 XX WPI; 2004-305150/28.  
 XX New polynucleotide sequence encoding multiple drug resistance proteins

PT from Drosophila melanogaster or Anopheles gambiae, useful in developing effective insecticides.  
 XX Claim 1; SEQ ID NO 6; 58pp; English.  
 CC The present invention relates to a purified polynucleotide or its fragment and comprises a sequence encoding multiple drug resistance CC proteins (MRPs) from Drosophila melanogaster or Anopheles gambiae. The polynucleotide is useful in developing effective insecticides. The CC present sequence represents human MRP1.  
 XX Sequence 1531 AA;  
 SQ  
 Query Match 78.5%; Score 7860; DB 8; Length 1531;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MALRGFCSADGSDPLMDWNTVNTSNPDFTKCFQNTVLVWVPCFVLMACFPFFYFLYLSRH 60  
 DB 1 MALRGFCSADGSDPLMDWNTVNTSNPDFTKCFQNTVLVWVPCFVLMACFPFFYFLYLSRH 60  
 QY 61 DRGVIQMTPLNKTKTALGFLWIVCWADLFYSFWERSRGIIFLAPVFLVSPILLGTTLLA 120  
 DB 61 DRGVIQMTPLNKTKTALGFLWIVCWADLFYSFWERSRGIIFLAPVFLVSPILLGTTLLA 120  
 QY 121 TFLIQLERRKGVQSSGIMLTFLVALVCALAILRSKIMTALKEDAQVDFRDITFYVYFS 180  
 DB 121 TFLIQLERRKGVQSSGIMLTFLVALVCALAILRSKIMTALKEDAQVDFRDITFYVYFS 180  
 QY 181 LLLIQLVLSGSDRSPLESETIHDNPNCPSSASFLSRITFWITGLIVRGYRQPLEGSD 240  
 DB 181 LLLIQLVLSGSDRSPLESETIHDNPNCPSSASFLSRITFWITGLIVRGYRQPLEGSD 240  
 QY 241 LWSLNKEDTSQVVPVLVKNWKKCAKTRKOPKVYSSKDPAPCKESSKVDANEVEAL 300  
 DB 241 LWSLNKEDTSQVVPVLVKNWKKCAKTRKOPKVYSSKDPAPCKESSKVDANEVEAL 300  
 QY 301 IVKSPQKEWNPFLFKVLYKTFPGYFLMSFFPKAIHDLMMFSGPOLKLLIKFVNDTKAPD 360  
 DB 301 IVKSPQKEWNPFLFKVLYKTFPGYFLMSFFPKAIHDLMMFSGPOLKLLIKFVNDTKAPD 360  
 QY 361 WQGYFYTVLLFVTACLOTILVHOYFHICFVSGMRIKTAIVGAVRKALVITNSARKSTV 420  
 DB 361 WQGYFYTVLLFVTACLOTILVHOYFHICFVSGMRIKTAIVGAVRKALVITNSARKSTV 420  
 QY 421 GEIVNLMSVDAQRFMDLATYINMLWSAPLOVILALYLNLGLPSVLGAVVWMLMVPV 480  
 DB 421 GEIVNLMSVDAQRFMDLATYINMLWSAPLOVILALYLNLGLPSVLGAVVWMLMVPV 480  
 QY 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAWELAFKDKVLAIHQEELKVLK 540  
 DB 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAWELAFKDKVLAIHQEELKVLK 540  
 QY 541 KSAVLSAVGTFTVCTPPLVALCTFAVVTVDENNILDAQTAFAVSLALFNILRFPNLILP 600  
 DB 541 KSAVLSAVGTFTVCTPPLVALCTFAVVTVDENNILDAQTAFAVSLALFNILRFPNLILP 600  
 QY 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWARSDDPT 660  
 DB 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWARSDDPT 660  
 QY 661 LNGTHSIPEGALVAVVGQVCGKSSLLSALLAEMDKVEGHVAIKGSVAYVPPQAWIQND 720  
 DB 661 LNGTHSIPEGALVAVVGQVCGKSSLLSALLAEMDKVEGHVAIKGSVAYVPPQAWIQND 720  
 QY 721 SLRENILFGCOLEBPYRYSVIOACALLPDLBILPSGDRTEIGEKVNLISGQKQVSLAR 780  
 DB 721 SLRENILFGCOLEBPYRYSVIOACALLPDLBILPSGDRTEIGEKVNLISGQKQVSLAR 780  
 QY 781 AVYNADIYLFDDPLSAVDHVGKHIPENVIGPKMLKNKTRILVTHSMYSYLPQVDVIV 840  
 DB 781 AVYNADIYLFDDPLSAVDHVGKHIPENVIGPKMLKNKTRILVTHSMYSYLPQVDVIV 840

QY 841 MSGGKISMGYSOELLARDGAFELRTYASTQEODAEENGVTGYSGPCKEAKOMENG 900  
 DB 841 MSGGKISMGYSOELLARDGAFELRTYASTQEODAEENGVTGYSGPCKEAKOMENG 900  
 QY 901 LVTDGAGQIQORQLSSSSSYSGDISRHNSHSTAELOKAEAKKEETWKLMEADKAQTQGVKL 960  
 DB 901 LVTDGAGQIQORQLSSSSSYSGDISRHNSHSTAELOKAEAKKEETWKLMEADKAQTQGVKL 960  
 QY 961 SVYDYMKAICLPIFSLFIFLFCNHNVSALASNYWLSLWTDPIVNGTQEHKTVRLSVYG 1020  
 DB 961 SVYDYMKAICLPIFSLFIFLFCNHNVSALASNYWLSLWTDPIVNGTQEHKTVRLSVYG 1020  
 QY 1021 ALGISQIAVFGYSMAVSGIGILASRLHVDLHLSILRSPMSFFERTPSGNLVNRSKEL 1080  
 DB 1021 ALGISQIAVFGYSMAVSGIGILASRLHVDLHLSILRSPMSFFERTPSGNLVNRSKEL 1080  
 QY 1081 DTVDSDIPEVIMFMGSLFNIVGACIVILLATPIAAIIIPPLGLIYFFVQRFVASSRQL 1140  
 DB 1081 DTVDSDIPEVIMFMGSLFNIVGACIVILLATPIAAIIIPPLGLIYFFVQRFVASSRQL 1140  
 QY 1141 KRLESVSRSPVYSHNETLLGVSVIRAFBEQERFIHQSDIKVDENOKAYPISIVANRWLA 1200  
 DB 1141 KRLESVSRSPVYSHNETLLGVSVIRAFBEQERFIHQSDIKVDENOKAYPISIVANRWLA 1200  
 QY 1201 VRLCVCNCIIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTVLNLVRMSSEMETNIVA 1260  
 DB 1201 VRLCVCNCIIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTVLNLVRMSSEMETNIVA 1260  
 QY 1261 VERLKEYSTEKAPQIQETAPPSSWPQVGRVFRNRYCLRYREDLDFVLRHINVTINGG 1320  
 DB 1261 VERLKEYSTEKAPQIQETAPPSSWPQVGRVFRNRYCLRYREDLDFVLRHINVTINGG 1320  
 QY 1321 EKVGIIVGRTGAGKSSLTGLFRINSAGEIIGDINIAKIGHDLRFKTIIPQDPVLF 1380  
 DB 1321 EKVGIIVGRTGAGKSSLTGLFRINSAGEIIGDINIAKIGHDLRFKTIIPQDPVLF 1380  
 QY 1381 SGLSRNLDPPFQYSDDEEVTWTSLELAHLKDFVSALPDKLDHECAGGENLSVGQRLVCL 1440  
 DB 1381 SGLSRNLDPPFQYSDDEEVTWTSLELAHLKDFVSALPDKLDHECAGGENLSVGQRLVCL 1440  
 QY 1441 ARALARCKTILVLDATAAVDLETDLQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 DB 1441 ARALARCKTILVLDATAAVDLETDLQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 QY 1501 DKGEIQEYGAPELDLQQRGLFYSMADAGLV 1531  
 DB 1501 DKGEIQEYGAPELDLQQRGLFYSMADAGLV 1531

RESULT 14

ADN95929  
 XX ADN95929 standard; protein; 1530 AA.  
 XX AC ADN95929;  
 XX DT 01-JUL-2004 (first entry)  
 XX DE Human BEC/LEC-related protein sequence SeqID853.  
 XX KW growth; differentiation; blood endothelial cell; BEC;  
 KW lymphatic endothelial cell; LEC; hereditary lymphoedema; VEGF-3;  
 KW lymphatic growth agent; VEGF-C; VEGF-D; antiangiogenic; cytostatic;  
 KW vasotrophic; antiinflammatory; gene therapy; endothelial cell disorder;  
 KW inflammatory disease; cancer metastasis; lymphatic system; human.  
 OS Homo sapiens.  
 XX WO2003080640-A1.  
 XX PD 02-OCT-2003.  
 XX PF 07-MAR-2003; 2003WO-US006900.  
 XX

PR 07-MAR-2002; 2002US-0363019P.  
 XX (LUDW-) LUDWIG INST CANCER RES.  
 PA (LICN ) LICENTIA LTD.  
 PA XX Alitalo K, Makinen T, Petrova T, Saharinen P, Saharinen J;  
 PI WPI; 2003-876899/81.  
 DR N-PSDB; ADN95930.  
 DR XX Example 1; SEQ ID NO 853; 176pp; English.  
 XX This invention relates to a method of differentially modulating the  
 CC growth or differentiation of blood endothelial cells (BEC) or lymphatic  
 CC endothelial cells (LEC) comprises contacting endothelial cells with a  
 CC composition comprising an agent that differentially modulates blood or  
 CC lymphatic endothelial cells. Treating hereditary lymphoedema comprises  
 CC identifying a human subject with lymphoedema and with a mutation in at  
 CC least one allele of a gene encoding a LEC protein, where the mutation  
 CC correlates with lymphoedema in human subjects, and with the proviso that  
 CC the LEC protein is not VEGF-3; and administering to the subject a  
 CC composition comprising a lymphatic growth agent selected from VEGF-C or  
 CC VEGF-D polypeptides and polynucleotides. The invention may be useful for  
 CC the development of compounds with an antiangiogenic, cytostatic,  
 CC vasotropic or antiinflammatory activity or for gene therapy. The method  
 CC is useful in modulating the growth or differentiation of blood  
 CC endothelial cells or lymphatic endothelial cells, in treating hereditary  
 CC lymphoedema, in screening for an endothelial cell disorder or  
 CC predisposition to the disorder or in monitoring the efficacy or toxicity  
 CC of a drug on endothelial cells. The agent is useful in manufacturing a  
 CC medicament for the differential modulation of blood vessel endothelial  
 CC cell or lymphatic vessel endothelial cell growth or differentiation. The  
 CC lymphatic growth agent may also be used in manufacturing a medicament for  
 CC the treatment of hereditary lymphoedema resulting from a mutation in a  
 CC LEC gene or of other diseases involving the lymphatic vessels, such as  
 CC various inflammatory diseases and cancer metastasis via the lymphatic  
 CC system. The present sequence is that of a human LEC/BEC differentially  
 CC expressed protein which is related to the method of the invention. Note:  
 CC This sequence does not appear in the specification but was obtained by  
 CC the indexer using the source data given in table 14 of the specification.  
 XX SQ Sequence 1530 AA;

Query Match 78.4%; Score 7856; DB 7; Length 1530;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1530; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MALRGFCADGSDPLWDNVTWNTSNPDFTKCFONTVLVWVPCFYLWACPPFFLYLSRH 60  
 DB 1 MALRGFCADGSDPLWDNVTWNTSNPDFTKCFONTVLVWVPCFYLWACPPFFLYLSRH 60  
 QY 61 DRGIQMTPLNKTALGFLLLWVWADLFYSFWSRGIPLAPVFLVSPITLLGTTLLA 120  
 DB 61 DRGIQMTPLNKTALGFLLLWVWADLFYSFWSRGIPLAPVFLVSPITLLGTTLLA 120  
 QY 121 TFLIQLERRRGVQSSGIMLTFVLVLCALILRSKIMTALKEDAQVLDLPRDITFYVYFS 180  
 DB 121 TFLIQLERRRGVQSSGIMLTFVLVLCALILRSKIMTALKEDAQVLDLPRDITFYVYFS 180  
 QY 181 LLLIQLVLSCFSDRSPLFSETIHDNPNCPSSASFLSRITFWITGLIVRGYRQPLEGSD 240  
 DB 181 LLLIQLVLSCFSDRSPLFSETIHDNPNCPSSASFLSRITFWITGLIVRGYRQPLEGSD 240  
 QY 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRQPKVYVSSKDPAPQKSSKVDANEEVEAL 300  
 DB 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRQPKVYVSSKDPAPQKSSKVDANEEVEAL 300  
 QY 301 IVKSPQKWNPSLFKVLKTYTFFGYFLMSPPFKAIHDLMMFSGPQILKLLKFNVDTKAPD 360  
 DB 301 IVKSPQKWNPSLFKVLKTYTFFGYFLMSPPFKAIHDLMMFSGPQILKLLKFNVDTKAPD 360  
 QY 361 WQGYFTVLLFVTACIQTLVHLHOYFHCVFVSGMNRKIKTAVIGAYVRKALVTNSARKSSTV 420

Db 361 WQGYFTVLLFTVACIQLTVLHQYFHI CFVSGNRKIATVIGAVYRKALVITNSARKSTV 420  
Qy 421 GEIVNLMSVDAQRFMDLATYINNIWSAPLQVILALYLLMLNLGSPSVLAGVAVVLMVFN 480  
Db 421 GEIVNLMSVDAQRFMDLATYINNIWSAPLQVILALYLLMLNLGSPSVLAGVAVVLMVFN 480  
Qy 481 AVAMKTKTYQVAHMKSKNRIKLMEIINGIKVLKLYAWELAFKDKVLAIRQELKVLK 540  
Db 481 AVAMKTKTYQVAHMKSKNRIKLMEIINGIKVLKLYAWELAFKDKVLAIRQELKVLK 540  
Qy 541 KSAYLSAVCTFTWCTPFLVALCTFAVYVITDENNILDQAOTAFVSLALENIRFPLNLLP 600  
Db 541 KSAYLSAVCTFTWCTPFLVALCTFAVYVITDENNILDQAOTAFVSLALENIRFPLNLLP 600  
Qy 601 WVISSIVQASVSLKRLRIFLSHEELEPDSIERPPVKDGGGTNSITVRNATFTWASDPT 660  
Db 601 WVISSIVQASVSLKRLRIFLSHEELEPDSIERPPVKDGGGTNSITVRNATFTWASDPT 660  
Qy 661 LINGITPSIPEGALVAVVGQVGGCKSSLLSALLAEMDKVEHVAIKGSVAVYVQQAQWIND 720  
Db 661 LINGITPSIPEGALVAVVGQVGGCKSSLLSALLAEMDKVEHVAIKGSVAVYVQQAQWIND 720  
Qy 721 SURENLLFCQLEEPYRYSVIOACALLPDLILPSGDRTEIGCKGNLSGGQKQVSLAR 780  
Db 721 SURENLLFCQLEEPYRYSVIOACALLPDLILPSGDRTEIGCKGNLSGGQKQVSLAR 780  
Qy 781 AVYSNADIYLFDDPLSAVDHVGKHI FENVIGPKMLKNKTRILVTHSMSYLPQVDVIV 840  
Db 781 AVYSNADIYLFDDPLSAVDHVGKHI FENVIGPKMLKNKTRILVTHSMSYLPQVDVIV 840  
Qy 841 MSGGKISEMGVSOELLAROGAFABFLRTYASTEQODAEENGVTGSGPGKEAKQWENG 900  
Db 841 MSGGKISEMGVSOELLAROGAFABFLRTYASTEQODAEENGVTGSGPGKEAKQWENG 900  
Qy 901 LVTDSAGLQRLQSSSSSYSGDISRHNSSTAELOKAEAKKEETWKLMEADKAQGVKL 960  
Db 901 LVTDSAGLQRLQSSSSSYSGDISRHNSSTAELOKAEAKKEETWKLMEADKAQGVKL 960  
Qy 961 SVYWDYMKAIGLFISFLSIFLPMCHVNSALSNYLSLWTDPIVNGTOEHTKVLRSVYG 1020  
Db 961 SVYWDYMKAIGLFISFLSIFLPMCHVNSALSNYLSLWTDPIVNGTOEHTKVLRSVYG 1020  
Qy 1021 ALGISQGIADVGSMAVSTGGILASRCLHVDLHLSILRSPMSFFERTPSGNLVNRFSKEL 1080  
Db 1021 ALGISQGIADVGSMAVSTGGILASRCLHVDLHLSILRSPMSFFERTPSGNLVNRFSKEL 1080  
Qy 1081 DTVDSPMPVIMKFMGSLFNVIACIVILLATPIAAIIIPPLGLIYFFVQRYFVASSRQL 1140  
Db 1081 DTVDSPMPVIMKFMGSLFNVIACIVILLATPIAAIIIPPLGLIYFFVQRYFVASSRQL 1140  
Qy 1141 KRLSVSRSPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENOKAYYPSIVANRWLA 1200  
Db 1141 KRLSVSRSPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENOKAYYPSIVANRWLA 1200  
Qy 1201 VRLCVCNGLVILFAALFAVISRHSLSAGLVGLSVSLQVTTYLNLWRMSSEMETNIVA 1260  
Db 1201 VRLCVCNGLVILFAALFAVISRHSLSAGLVGLSVSLQVTTYLNLWRMSSEMETNIVA 1260  
Qy 1261 VERLKEYSTEKAPWQIETAPPSPQVGRVFRNRYCLIRYREDLDLFLVRHINTYINGG 1320  
Db 1261 VERLKEYSTEKAPWQIETAPPSPQVGRVFRNRYCLIRYREDLDLFLVRHINTYINGG 1320  
Qy 1321 EKVGVGRTGAGKSSLTGLFRINSAGEEIIIDGINIAKIGHDLRFKTIIPDPVLF 1380  
Db 1321 EKVGVGRTGAGKSSLTGLFRINSAGEEIIIDGINIAKIGHDLRFKTIIPDPVLF 1380  
Qy 1381 SGLSRMLNLPFSQYDEEYVTSLELAHLKDFYSALPDKLDHECAGGENLSVGQRLVCL 1440  
Db 1381 SGLSRMLNLPFSQYDEEYVTSLELAHLKDFYSALPDKLDHECAGGENLSVGQRLVCL 1440  
Qy 1441 ABALLARKILVLDENATAVDLETDDLIOSTTRTQFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
Db 1441 ABALLARKILVLDENATAVDLETDDLIOSTTRTQFEDCTVLTIAHRLNTIMDYTRVIVL 1500

Qy 1501 DKGEIQEYCAPSDLLQQRGLFYSMKADAGL 1530  
Db 1501 DKGEIQEYCAPSDLLQQRGLFYSMKADAGL 1530

## RESULT 15

AAR54928  
ID AAR54928 standard; protein; 1531 AA.

XX AAR54928;

XX 25-MAR-2003 (revised)  
DT 14-OCT-1994 (first entry)

XX Multidrug resistance protein.

XX Multidrug resistance protein; MRP; H69AR; cancer cell line; stem cell;  
KW cardiac muscle; transgenic animal.

XX Homo sapiens.

XX WO9410303-A1.

XX 11-MAY-1994.

XX 27-OCT-1993; 93WO-CA000439.

XX 27-OCT-1992; 92US-00966923.

XX 08-MAR-1993; 93US-00029340.

XX (TOOH ) UNIV QUEBENS KINGSTON.

XX Deeley RG, Cole SPC;

XX WPI; 1994-167460/20.

XX N-PSDB; AAQ65377.

XX Multi-drug resistance gene - encodes protein capable of conferring multi-  
drug resistance on cells, useful in diagnostic and treatment methods.

XX Disclosure; Page 69-74; 101pp; English.

XX The multidrug resistant cancer cell line H69AR (ATCC CRL 11350) was used  
to identify cDNA encoding a novel protein associated with multidrug  
resistance, MRP. MRP may be expressed in e.g. hematopoietic stem cells or  
cardiac muscle, or in transgenic animals, or can be used to raise  
antibodies. (Updated on 25-MAR-2003 to correct FN field.)

XX Sequence 1531 AA;

Query Match 78.4%; Score 7849; DB 2; Length 1531;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MALRGFCADSGSDPLMDNNTWNTSNPDKTQFNTLVWVPCFYLWACFPFFLYLSRH 60  
Db 1 MALRGFCADSGSDPLMDNNTWNTSNPDKTQFNTLVWVPCFYLWACFPFFLYLSRH 60  
Qy 61 DRGYIQTMTPLNKTALGFLWLWVCWADLFYSFMRSGIFLAPVFLVSPDLLGITTLA 120  
Db 61 DRGYIQTMTPLNKTALGFLWLWVCWADLFYSFMRSGIFLAPVFLVSPDLLGITTLA 120  
Qy 121 TFLQLERKGVQSSGIMLTFWLVVALCALILRSKIMTALKEDAQVDLFRDITFVYVFS 180  
Db 121 TFLQLERKGVQSSGIMLTFWLVVALCALILRSKIMTALKEDAQVDLFRDITFVYVFS 180  
Qy 181 LLLIQLVLSGFSRSPLESETIHDNPNCPSSASFLSRITFWITGLIVRGVQPLEGSD 240  
Db 181 LLLIQLVLSGFSRSPLESETIHDNPNCPSSASFLSRITFWITGLIVRGVQPLEGSD 240  
Qy 241 LWSLNKEDTSQVVPVLVKNWKKCAKTRKQPVKVVYSSKDPAPKSSKVDANESEVAL 300  
Db 241 LWSLNKEDTSQVVPVLVKNWKKCAKTRKQPVKVVYSSKDPAPKSSKVDANESEVAL 300

Db 241 LWSLNKEDTSEQVVPVVLVKNWKECAKTRKQPVKVYVSSKDPQPKESSKVDANEVEAL 300  
 QY 301 IVKSPKQWNPFLFKVLYKTGPFYFLMSFFKAIHDLMPFSGPOIILKLIKFNVDTKAPD 360  
 Db 301 IVKSPKQWNPFLFKVLYKTGPFYFLMSFFKAIHDLMPFSGPOIILKLIKFNVDTKAPD 360  
 QY 361 WQGYFYTVLLFVTFACLTQLVHLQYFHCYVSGMRKTAIVIGAVYRKALVITNSARKSSTV 420  
 Db 361 WQGYFYTVLLFVTFACLTQLVHLQYFHCYVSGMRKTAIVIGAVYRKALVITNSARKSSTV 420  
 QY 421 GEIVNLMSVDAQRFMDLAIYINMIWSAPLOVIALYLLMLNIGPSVLAVVWLMVVPV 480  
 Db 421 GEIVNLMSVDAQRFMDLAIYINMIWSAPLOVIALYLLMLNIGPSVLAVVWLMVVPV 480  
 QY 481 AVMAKTKTYQVAHMKSKONRIKMEIILNGIKVLYAWELAFKDKVLAIQEBELKVLK 540  
 Db 481 AVMAKTKTYQVAHMKSKONRIKMEIILNGIKVLYAWELAFKDKVLAIQEBELKVLK 540  
 QY 541 KSAYLSAVGTFTWVCTPFLVALCTFAVYVTTIDENNILDAQTAFVSLAFNLIRFPLNILP 600  
 Db 541 KSAYLSAVGTFTWVCTPFLVALCTFAVYVTTIDENNILDAQTAFVSLAFNLIRFPLNILP 600  
 QY 601 WYISSIVQASVSLKRLRIFLSHHELEPDSIERRPVKDGGTNSITVRNATFTWASDPPT 660  
 Db 601 WYISSIVQASVSLKRLRIFLSHHELEPDSIERRPVKDGGTNSITVRNATFTWASDPPT 660  
 QY 661 LINGITFPIPEGALVAVVGQVCGKSLLSALLAEMDKVEGHVAIKGSVAYVPPQAWIOND 720  
 Db 661 LINGITFPIPEGALVAVVGQVCGKSLLSALLAEMDKVEGHVAIKGSVAYVPPQAWIOND 720  
 QY 721 SURENILFCQLEEPYRSVIOACALLPDLLEILPSGDRTEIGEKGVNLSGGQKQVSLAR 780  
 Db 721 SURENILFCQLEEPYRSVIOACALLPDLLEILPSGDRTEIGEKGVNLSGGQKQVSLAR 780  
 QY 781 AVYSNADIYLFDDPLSADVAHVCKHI FENVIGPKMLKNKTRILVTHSMSYLPQVDVIV 840  
 Db 781 AVYSNADIYLFDDPLSADVAHVCKHI FENVIGPKMLKNKTRILVTHSMSYLPQVDVIV 840  
 QY 841 MSGGKISEMGSYQELLARDGAFELRTYASTEQDQDAEENGVTGVSFGPKAKOMENG 900  
 Db 841 MSGGKISEMGSYQELLARDGAFELRTYASTEQDQDAEENGVTGVSFGPKAKOMENG 900  
 QY 901 LVTDSAGLQRLQSSSSSYSGDISRHNSHSTAELQAKBAKKEETWKLMEADKAQTGVKL 960  
 Db 901 LVTDSAGLQRLQSSSSSYSGDISRHNSHSTAELQAKBAKKEETWKLMEADKAQTGVKL 960  
 QY 961 SVYDYMKAIGLFI SFLIFLPMCNHVSALASNYWLSLWTDPIVNGTQHTKVRLSVYG 1020  
 Db 961 SVYDYMKAIGLFI SFLIFLPMCNHVSALASNYWLSLWTDPIVNGTQHTKVRLSVYG 1020  
 QY 1021 ALGISQGI AVFGYSMAVSGIGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFSKEL 1080  
 Db 1021 ALGISQGI AVFGYSMAVSGIGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFSKEL 1080  
 QY 1081 DTVDSMIPEVIMKFMGSLFNIGACIVILLATPIAAIIPPLGLIYFFVQRFYVASSRQL 1140  
 Db 1081 DTVDSMIPEVIMKFMGSLFNIGACIVILLATPIAAIIPPLGLIYFFVQRFYVASSRQL 1140  
 QY 1141 KRLESVSRSPVYSHENETLLGVSVIRAFEEQERFIHQSDLKVDENOKAYPSIVANRWLA 1200  
 Db 1141 KRLESVSRSPVYSHENETLLGVSVIRAFEEQERFIHQSDLKVDENOKAYPSIVANRWLA 1200  
 QY 1201 VRLECVGNICVILFAALFAVISRHSLSAGLSVSYSLQVTTYLNWLVMSSEMETNIVA 1260  
 Db 1201 VRLECVGNICVILFAALFAVISRHSLSAGLSVSYSLQVTTYLNWLVMSSEMETNIVA 1260  
 QY 1261 VERLKEYSETKEAPWQIQTETAPSSWPQGRVFRNYCLRYREDLPVLRHINVTINGG 1320  
 Db 1261 VERLKEYSETKEAPWQIQTETAPSSWPQGRVFRNYCLRYREDLPVLRHINVTINGG 1320  
 QY 1321 EKVGVGRTGAGKSSLTGLFRINESAECEI IIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
 Db 1321 EKVGVGRTGAGKSSLTGLFRINESAECEI IIDGINIAKIGLHDLRFKTIIPQDPVLF 1380

QY 1381 SGSLRMNLDPPFSQYSDSEVWTSLELAHLKDFVSALPDKLDHECAGGENLSVGQRQLVCL 1440  
 Db 1381 SGSLRMNLDPPFSQYSDSEVWTSLELAHLKDFVSALPDKLDHECAGGENLSVGQRQLVCL 1440  
 QY 1441 ARALLRKTKILVLDATAAVALDETDDLQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 Db 1441 ARALLRKTKILVLDATAAVALDETDDLQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 QY 1501 DKGETOEYCAPSDLLQQRGLFYSMADAGLV 1531  
 Db 1501 DKGETOEYCAPSDLLQQRGLFYSMADAGLV 1531

Search completed: March 18, 2005, 11:02:20  
 Job time : 146.718 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2005, 10:57:20 ; Search time 45.7349 seconds  
(without alignments)  
4096.085 Million cell updates/sec

Title: US-10-665-283-8  
Perfect score: 10016  
Sequence: 1 MALRGFCSADGSDPLMDWNV.....RSVAVAKPKFSISPDLS 1947

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	7849	78.4	1531	1	DVHUAR
2	4484.5	44.8	1527	2	multidrug resistan
3	3587	35.8	1545	1	canalicular multia
4	3507.5	35.0	1541	1	multidrug resistan
5	3289	32.8	1494	2	canalicular multia
6	3225	32.2	1573	2	protein F57C12.4 (
7	3195.5	31.9	1502	2	hypothetical prote
8	2838	28.3	1515	1	multidrug resistan
9	2732	27.3	1398	2	cadmium resistance
10	2548	25.4	1478	2	hypothetical prote
11	2384.5	23.8	1623	2	ABC transporter SP
12	2356.5	23.5	1622	2	ABC transporter At
13	2326	23.2	1495	2	glutathione S-conj
14	2312	23.1	1144	2	probable ABC trans
15	2294.5	22.9	1559	1	hypothetical prote
16	2263.5	22.6	1488	2	probable ABC trans
17	2238.5	22.3	1516	2	glutathione-conjug
18	2230	22.3	1539	2	ABC transporter-11
19	2144.5	21.4	1355	2	hypothetical prote
20	2144.5	21.4	1514	2	multi resistance p
21	2140	21.4	1515	2	MRP-like ABC trans
22	2131.5	21.3	1490	2	multi resistance p
23	2129	21.3	1545	2	sulfonylurea recep
24	2123	21.2	1153	2	hypothetical prote
25	2109	21.1	1545	2	sulfonylurea recep
26	2104.5	21.0	1511	2	sulfonylurea recep
27	2098	20.9	1546	2	sulfonylurea recep
28	2069.5	20.7	1389	2	ABC transporter-11
29	2062.5	20.6	1661	2	probable membrane

ALIGNMENTS

RESULT 1

DVHUAR

multidrug resistance protein (cell line H69AR) - human  
N:Alternate names: multidrug resistance-associated protein (MRP)

C:Species: Homo sapiens (man)

C>Date: 30-Jun-1993 #sequence\_revision 05-Dec-1998 #text\_change 13-Jan-2001

C:Accession: A44231; A37495

R:Coile, S.P.C.; Bhardwaj, G.; Gerlach, J.H.; Mackie, J.E.; Grant, C.E.; Almquist, K.C.;

Science 258, 1650-1654, 1992

A:Title: Overexpression of a transporter gene in a multidrug-resistant human lung cancer.

A:Reference number: A44231; MUID:93088080; PMID:1360704

A:Accession: A44231

A>Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: MAPRTGTGMSRGIPATPTSPAFRRSSCGLVFTSGPV, 50-1531 <Col>

A:Cross-references: GB:L05628; NID:gl835658

A:Experimental source: small cell lung carcinoma cell line H69AR

A>Note: sequence extracted from NCBI backbone (NCBIP:119851); this sequence has been co

R:Coile, S.P.C.; Deeley, R.G.

Science 260, 879, 1993

A:Title: Multidrug resistance-associated protein: sequence correction.

A:Reference number: A37495; MUID:93262415; PMID:8098549

A:Accession: A37495

A>Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-60 <CO2>

A:Cross-references: GB:L05628; NID:gl835658

A>Note: sequence extracted from NCBI backbone (NCBIP:131929)

C:Genetics:

A:Gene: GDB:MRP

A:Cross-references: GDB:136335; OMIM:158343

A:Map position: 16p13.1-16p13.1

C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology

C:Keywords: antibiotic resistance; ATP; duplication; nucleotide binding; P-loop; transp

F:661-844/Domain: ATP-binding cassette homology <ABC1>

F:678-685/Region: nucleotide-binding motif A (P-loop)

F:788-792/Region: nucleotide-binding motif B

F:1310-1503/Domain: ATP-binding cassette homology <ABC2>

F:1327-1334/Region: nucleotide-binding motif A (P-loop)

F:1450-1454/Region: nucleotide-binding motif B

Query Match 78.4%; Score 7849; DB 1; Length 1531;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MALRGFCSADGSDPLMDWNVNTNTSNPDTKCFQNTLVWVPCFYLWACFPFYLSRH 60

Db 1 MALRGFCSADGSDPLMDWNVNTNTSNPDTKCFQNTLVWVPCFYLWACFPFYLSRH 60

QY 61 DRGYIQMTPLNKTALGFLLWIVCWADLFYFWSRSRGIFLAPFLVSPFTLLGITLLA 120

Db 61 DRGYIQMTPLNKTALGFLLWIVCWADLFYFWSRSRGIFLAPFLVSPFTLLGITLLA 120

QY 121 TFLIQLERRKGVSQSGIMLTFWLVALYCALAIRSKIMTALKEDAQVDLFRDITFYVYFS 180  
 DB 121 TFLIQLERRKGVSQSGIMLTFWLVALYCALAIRSKIMTALKEDAQVDLFRDITFYVYFS 180  
 QY 181 LLLIQLVLSCFSDRSPLFSTIHDNPNCPSSASFLSRITFWITGLIVRGYRQPLEGSD 240  
 DB 181 LLLIQLVLSCFSDRSPLFSTIHDNPNCPSSASFLSRITFWITGLIVRGYRQPLEGSD 240  
 QY 241 LWSLNKEDTSEQVVPVVLVKNWKECAKTRKQPVKVYVSSKDPAPQPKSSKVDANEVEAL 300  
 DB 241 LWSLNKEDTSEQVVPVVLVKNWKECAKTRKQPVKVYVSSKDPAPQPKSSKVDANEVEAL 300  
 QY 301 IVKSPQKEMNPSPFLKVLTKTGPVFLMSPPFKAIHDLMMFSGPQIILKLIKLVNDTKAPD 360  
 DB 301 IVKSPQKEMNPSPFLKVLTKTGPVFLMSPPFKAIHDLMMFSGPQIILKLIKLVNDTKAPD 360  
 QY 361 WQGYFTVLLFVTFACLTQLVLHQYFHCIFVSGMRITKAVIGAYYRKALVITNSARKSSTV 420  
 DB 361 WQGYFTVLLFVTFACLTQLVLHQYFHCIFVSGMRITKAVIGAYYRKALVITNSARKSSTV 420  
 QY 421 GEIVNLSMVDQAQFMDLATVINNIWSAPLOVILALYLLNLGSPSVLAGVAVNVLMPVN 480  
 DB 421 GEIVNLSMVDQAQFMDLATVINNIWSAPLOVILALYLLNLGSPSVLAGVAVNVLMPVN 480  
 QY 481 AVNMAKTKTYQVAHMSKDNRIKLMNEILNGIKVLKLYAWELAFKDKVLAIQBEELKVLK 540  
 DB 481 AVNMAKTKTYQVAHMSKDNRIKLMNEILNGIKVLKLYAWELAFKDKVLAIQBEELKVLK 540  
 QY 541 KSAYLSAVGTFTWCTPFLVALCTPAVYVTDENNILDAQTAFVSLALFNILRFPNLILP 600  
 DB 541 KSAYLSAVGTFTWCTPFLVALCTPAVYVTDENNILDAQTAFVSLALFNILRFPNLILP 600  
 QY 601 MWISSIVQASVSLKRLRIFLSHEELPDSIERPVKDGGSITSITVRNATFWASDDPT 660  
 DB 601 MWISSIVQASVSLKRLRIFLSHEELPDSIERPVKDGGSITSITVRNATFWASDDPT 660  
 QY 661 LINGITSIPEGALVAVGVGCGKSLLSALLAEMDKVEGHVAIKGSVAYVPQAWIQND 720  
 DB 661 LINGITSIPEGALVAVGVGCGKSLLSALLAEMDKVEGHVAIKGSVAYVPQAWIQND 720  
 QY 721 SURENILFCQLEEPYRVSIVOCALLPLEILPSGDRTEIGKGVNLSGGQKQVSLAR 780  
 DB 721 SURENILFCQLEEPYRVSIVOCALLPLEILPSGDRTEIGKGVNLSGGQKQVSLAR 780  
 QY 781 AVYSNADIYLFDDPLSAVDHVGKHFENVIGPKGMLKNKTRILVTHSMVSLPQVDVLIIV 840  
 DB 781 AVYSNADIYLFDDPLSAVDHVGKHFENVIGPKGMLKNKTRILVTHSMVSLPQVDVLIIV 840  
 QY 841 MSGGKISEMGYSQELLARDGAPAEFLRTVASTEQDAAEENGVTGVSFGPKQKQWENG 900  
 DB 841 MSGGKISEMGYSQELLARDGAPAEFLRTVASTEQDAAEENGVTGVSFGPKQKQWENG 900  
 QY 901 LVYDSAGKQLQRLSSSSYSGDISRHNSSTAELQAKAEKEETWKLMEADKAQTQVKL 960  
 DB 901 LVYDSAGKQLQRLSSSSYSGDISRHNSSTAELQAKAEKEETWKLMEADKAQTQVKL 960  
 QY 961 SVYDYMKAIGLIFISFLIFPMCNHVSALASNYLSLWTDPIVNGTQEHTKVRLSVYG 1020  
 DB 961 SVYDYMKAIGLIFISFLIFPMCNHVSALASNYLSLWTDPIVNGTQEHTKVRLSVYG 1020  
 QY 1021 ALGISQGLAVFGYSMAVIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFSKEL 1080  
 DB 1021 ALGISQGLAVFGYSMAVIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFSKEL 1080  
 QY 1081 DTVDMSIEVIMKFMGSLFNIVIGACIVILLATPAAIIIPPLGLIYFPVQRYVASSRQL 1140  
 DB 1081 DTVDMSIEVIMKFMGSLFNIVIGACIVILLATPAAIIIPPLGLIYFPVQRYVASSRQL 1140  
 QY 1141 KRLESVSRSPVYSHNETLLGVSVIRAFEBQERTHQSDLKVDENQKAYPSIVANRWLA 1200  
 DB 1141 KRLESVSRSPVYSHNETLLGVSVIRAFEBQERTHQSDLKVDENQKAYPSIVANRWLA 1200

QY 1201 VRLECVGNCIVLFAALFAVSRHSLSAGLVGLSVSYSLQVTTYLNLVRMSSEMETNIVA 1260  
 DB 1201 VRLECVGNCIVLFAALFAVSRHSLSAGLVGLSVSYSLQVTTYLNLVRMSSEMETNIVA 1260  
 QY 1261 VERLKEYSSTEKEAPMOIQETAPPSSWPQVGRVERPNYCLRYREDLDFVLRHNVTINGG 1320  
 DB 1261 VERLKEYSSTEKEAPMOIQETAPPSSWPQVGRVERPNYCLRYREDLDFVLRHNVTINGG 1320  
 QY 1321 EKVGIIVGRTGACKSSITLGLFRINSAEGEIIGDINIAGIKGLHDLRFKTIIPQDPVLP 1380  
 DB 1321 EKVGIIVGRTGACKSSITLGLFRINSAEGEIIGDINIAGIKGLHDLRFKTIIPQDPVLP 1380  
 QY 1381 SCSLRNLDLPFQYSDDEEVTSLAHLKDFVSALPDKLDHCAEGENLSVGQRLVCL 1440  
 DB 1381 SCSLRNLDLPFQYSDDEEVTSLAHLKDFVSALPDKLDHCAEGENLSVGQRLVCL 1440  
 QY 1441 ABALLARKTKILVLDATAVLETDLQSTIRTFQEDCTVLTIAHRLNTIMDYTRVIL 1500  
 DB 1441 ABALLARKTKILVLDATAVLETDLQSTIRTFQEDCTVLTIAHRLNTIMDYTRVIL 1500  
 QY 1501 DKGETQYEGAPSDLLQORGLFYSMADAGLV 1531  
 DB 1501 DKGETQYEGAPSDLLQORGLFYSMADAGLV 1531  
 RESULT 2  
 JEO336  
 C:Species: Homo sapiens (man)  
 C:Date: 05-Feb-1999 #sequence\_revision 05-Feb-1999 #text\_change 09-Jul-2004  
 C:Accession: JEO336  
 R:Uchiumi, T.; Hinochita, E.; Haga, S.; Nakamura, T.; Tanaka, T.; Toh, S.; Furukawa, M.;  
 Biochem. Biophys. Res. Commun. 252, 103-110, 1998  
 A:Title: Isolation of a novel human canalicular multispecific organic anion transporter,  
 t.  
 A:Reference number: JEO336; MUID:99032812; PMID:9813153  
 A:Accession: JEO336  
 A>Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-1527 <UCH>  
 A:Cross-references: UNIPROT:O15438; GB:AF083552  
 C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology  
 C:Keywords: ATP  
 F:1306-1499/Domain: ATP-binding cassette homology <ABC2>  
 Query Match 44.8%; Score 4484.5; DB 2; Length 1527;  
 Best Local Similarity 56.6%; Pred. No. 3.2e-293;  
 Matches 871; Conservative 261; Mismatches 372; Indels 35; Gaps 9;  
 QY 8 SADGSDPLDWMNVTWNTSNPDFTKCFQNTVLVWVPCFVLWACFPFVFLYLSRHDGXYIQM 67  
 DB 7 SGEIGSKFWDNSLVHTENPDLTFCFQNSLAWVPCYLVWVALPCYLLYLRRHRCGYIIL 66  
 QY 68 TPLNKTATGLFLWIVCWADLFYSFWERSRGIFLAPVFLVSPTLGTLITLLATFLIQLE 127  
 DB 67 SHLSKLMWLGVLLWCVSADLFYSFHLVHGRAPAFVFFVPLVGVVTMLATLLIQYE 126  
 QY 128 RRKGVQSGGIMLTFWLVALYCALAIRSKIMTALKEDAQVDLFRDITFYVYFSLILQLV 187  
 DB 127 RLQGVQSGGVLIIIFWLCVCAIVPFPSKILLAKAEGEISDPFRFTFYTHFALVLSALI 186  
 QY 188 LSCFSDRSPLFSETIHDNPNCPSSASFLSRITFWITGLIVRGYRQPLEGSDLSLNKE 247  
 DB 187 LACFRKPPFPFSAKNVDNPNYPETSAGLSRLFPFWTKMAYIGYRHPLEKDLMSLKEE 246  
 QY 248 DTSEQVVPVVLVKNWKECAKTRKQPVKVYVSSKDPAPQPKSSKVDANEVEALIVKSPQK 307  
 DB 247 DRSQMVVQQLLEAWRQKQKTARH-----KASAPGK---NASGEDEVILGARPRP 294  
 QY 308 EWNPSLFLKLYKTFGPGVFLMSFFFKAIHDLMMFSGPQIILKLIKLVNDTKAPDQGVYFT 367  
 DB 295 R-NPSFLKALLAIFGSGFLISACFKLIQDLISLFINPQLLSILIRFISNPMGPMGWGFLVA 353





Db 344 GFVSSNSYWFYGCIAILMFVAVTLTQSFCLQSYFQHCFLVGCVRVTTVMSSYIKKALT 403  
 Qy 411 TNSARKSVTVGINLMSYDAQFMDLATYINNIKNSAPLOVTLALYLWLNGLSPVLAV 470  
 Db 404 SNLARKQYITIGFTVNLMSVDSQKMDATYMLQWSSVQITLSIFFLRELGPILLAV 463  
 Qy 471 AVMLVPVNAVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYLAWELAPKDKVLA 530  
 Db 464 GVMVLLIPVNGVLATKIRNIQVNMKNKDKRLKIMNEILSGIKLYFAWEPSPQEQVG 523  
 Qy 531 IQBELKVLKXSAIYLSAVCTFTWCTPFLVALCTFAVYVITIDENNILDQATFVSLALFN 590  
 Db 524 IRKELKNLRFQQLSFLIPTLOITPILVSVTFVSVYVLDVSDANVNAEKAFITILFN 583  
 Qy 591 ILRFPNLIPMWISSIVQASVSLKRLRIFLSHHELEPDSIERPRPVKGGGTSITVRNAT 650  
 Db 584 ILRFPNLMPTWSSILQASVSDRLERYLGGDDLDTSAIRVSNFD----KAVFSPAS 639  
 Qy 651 FTWASDPPTLNGITPISPEGALVAVGVGGCKSSLSALLAEMDKVEGHVAIKGSVAY 710  
 Db 640 FTWDPLEATIQVNDLIDKPGQLVAVGVGVGKSSLSVAMGEMENVHGHITIQSTAY 699  
 Qy 711 VPOQAWIQNDSRENILFCQLEBPYRSVIOACALLPDLEILPBGDRTEICEKGNVLSG 770  
 Db 700 VPOQSWIQNGTIKDNILFSEYNEKYYQVLRKACALLPDLEILPBGDMAEIGKGINLSG 759  
 Qy 771 GOKQVSLARAVYSNADYFLDDPLSADVAHVGHIFENVIGPGLMKLNKTRILYTHMS 830  
 Db 760 GOKQVSLARAYQADYIYLDPLSADVAHVGHIFENVIGPGLMKLNKTRILYTHGH 819  
 Qy 831 YLPQVDVITVMSGGKISEMGYSQELLARDGAFELRTYAS-TEQEQDAEENGVTGVSGP 889  
 Db 820 FLPQVDEIVVLGKTIENGYSYRDLDDKGVFARNWKTFMKHSQPEGEATVNN----- 872  
 Qy 890 GKEAKQMGMLVT-----DSAG-----KOLQQLSSSSSYSGDISRHHSNABEQ-- 935  
 Db 873 DSEAEDDDGLIPTWEEIPEDAASLAMRRENSLRRTLSRSSSSRRRGSKLSNLSKINV 932  
 Qy 936 ---KAEAKKEETWKLMEADKATQGVKLSVYDYMKAIGLFTSLFPLFCMCHVYSALAS 992  
 Db 933 NVLKEKEVEGOKLKEFEVTEKVKFSYLUKYIQAVGWMSILFIILFYGLNUNVAFIS 992  
 Qy 993 NYWLSLWT--DDPIVNGT---QBHTKVLRSVYGALGISQIAVFGYSMAVSIIGILASRCL 1048  
 Db 993 NLWLSAWTSDNLNGTNSSSHDMRIGVFGALGLAQICLLISTLSIYACRNASKAL 1052  
 Qy 1049 HVDLHSILRSPMSFPTPSGNLVNRPFSKELDTVDSDMIPKVMGSLFNVIACIVI 1108  
 Db 1053 HQQLLTNLRAPMRFPDTPPTGRVNRVFSGDISTVDDLLPQTLRSWMMCFGIAGTLVMI 1112  
 Qy 1109 LLAATPIAIIIPGLGIYFFQRFYVASSROLKRESYRSRSPVYSHFNETLLGVSVIRAF 1168  
 Db 1113 CMATPVFAIIIPLSILYISQVFFVYATSRQLRSDSVTKSPYSHFSETVGLPIRAF 1172  
 Qy 1169 EQERFIHOSDLKVDENOKAYPSIVANRWLVRLECVGNCIVLPAALFAVIRSHSLAG 1228  
 Db 1173 EQQRFLAWNEQIDINOKCVFSWITSNRLAIRLELVNVLVFCALLVLYRKLTCGD 1232  
 Qy 1229 LVGLSVYSQVTTVNLNLMVMSMETNIVAVERLKEYSETEKAPMOIQETAPSSWP 1288  
 Db 1233 VVGFVLSNALNITOTLNLVLRMTSEATNIVAVERTSEYINVENEAPW-VTDKRPADWP 1291  
 Qy 1289 QYRVEFRNYCLRYREDLDFLRHINVTINGEKYIGVCTGAGKSSITLGIFRINESAE 1348  
 Db 1292 RHGEIQFNQYQRYPELDELVLKGTCTNICKSEKYGWGTGAGKSSLTNCLFRILESAG 1351  
 Qy 1349 GIIIDGINIAIGHDLRFKTIIPQDPVLFSGSLRMNLDPFQSVDSBEVWTSLELAHL 1408  
 Db 1352 GOIIDGIDVASIGHDLRERITIIIPQDPILFSGSLRMNLDLPFKYSDEEVRWALELAHL 1411  
 Qy 1409 KDFVSALPDKLDECAEGENISVGQRLVCLARALLRKTILVLDEATAAVDLTDDLI 1468

Db 1412 RSFVSGLOGLLSEVTEGGDNLSIQORQLCLGRVLRKSKILVLDEATAAVDLTDSIJ 1471  
 Qy 1469 OSTRTQFEDCTVLTIAHRLNTIMDYRVILVDKGEIOBYGAPSDLLQORGLFYMAKDA 1528  
 Db 1472 OTTIRKESQCTVITIAHRLTIMDSKIMVLDNGKIVEYGSPEELLSNRGSFYLMAKEA 1531  
 Qy 1529 GL 1530  
 Db 1532 GI 1533  
 RESULT 5  
 E89447  
 protein F57C12.4 [imported] - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Jul-2004  
 C:Accession: E89447  
 R:Anonymous, The C. elegans Sequencing Consortium.  
 Science 282, 2012-2018, 1998  
 A:Title: Genome sequence of the nematode C. elegans: a platform for investigating bio-  
 A:Reference number: A75000; PMID:9851916  
 A:Ref: see websites genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/Projects/C\_e-  
 A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and  
 A:Accession: E89447  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-1494 <STO>  
 A:Cross-references: UNIPROT:Q20943; GB:chr\_X; PIDN:AAA83299.1; PID:g1118071; GSPDB:GN00C  
 C:Map position: X  
 C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology  
 Query Match 32.8%; Score 3289; DB 2; Length 1494;  
 Best Local Similarity 47.6%; Pred. No. 1.1e-212;  
 Matches 696; Conservative 239; Mismatches 435; Indels 92; Gaps 18;  
 Qy 119 LATFLIQLERRKGVSSGIMLTFLVAVLVCALAILRSKIMTALKEDAQVDFRDTFFVY 178  
 Db 76 LALILTVACKNGIITSGVITLYWLVVCGIPEPRFVLSGFIYNEYALEGIRATLYIA 135  
 Qy 179 FSLLIQLVLCSPDRSPLFETHIDPNCPRESSASFLSRITFTFWITGLIVRGYQPLRG 238  
 Db 136 FTFSALELFCFAD---VPSDMYKSESSCPETASFINRLTFQWFTGLAYLGNKSKLEN 192  
 Qy 239 SDSLNLNKEDSEQVVPVLVKNWKKCAKTRQPKVYVYSSKDP--AOPKSSSKVDANE 296  
 Db 193 EDLMDLNEIDKAENLIPSMQNLKPRIDEYH-QNIK-----KPSAALPKN----- 237  
 Qy 237 VEALIVKSPQKWNPSLFKLYKTFPGPYFLMSFFPKAIHDLMMFSGPQILKLLIKFVNDT 356  
 Db 238 -----HPSFVIPFKTYKTYLLAGFFYKLCFCDMLQFLAPQLLKQLIGFIEDK 284  
 Qy 357 KAPDHQGYFVTVLLPVTACLOTVLVHQFHCFTVSGMRIKTAIVGAVTRKALVINRSARK 416  
 Db 285 NOPVMIGSIVGIMPFSSFLQSMFLHQYVHSMFRGMVRSVLTSAVYSKALNLSNEARK 344  
 Qy 417 SSTVGEIVNLSVDAQRFMDLATYINMTWSAPLOVILALYLWLNGLSPVLAVMLM 476  
 Db 345 GKTIGAVNLSVSDIQIADMPTIMLFWSPAPLQIFLSYIFLWKFGLVAALAGLVLLIA 404  
 Qy 477 VPVNAVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYLAWELAPKDKVLAIRQEBL 536  
 Db 405 LPVNGLIAIQKRCQTEQMKLKDBERIKWSEILNGMKVLYSWERSMENWVLIKIREREL 464  
 Qy 537 KVLKKSAYLSAVGTFTWCTPFLVALCTFAVYVTTID-ENNILDAQTAFVSLALFNILRPP 595  
 Db 465 HILKLSYFMAAIVPSWICAPFLASVIFVYVYLDPENNVLTPEITTFVALSFLDIRMP 524  
 Qy 596 LNILPMVYSSIVQASVSLKRLRIFLSHHELEPDSIERPRPVKGGGTSITVRNATFTWAR 655  
 Db 525 LAMVAVYGEAVQCVSVNTRLKEFFAAEMSPQT-----SISHGETSDSAIEVNGULFSWSS 580





Db 887 -RPRSDAAPVKGS-----TSEAQMPEPLDDVEVTGLTAGEDSGVQYGRVKSAFYLSVLA 940

QY 970 IGLFISFLSLFMCNHSALSNYLSLWTDPIVNGTQEHFKVRLSVYGALIGSQGTA 1029

Db 941 VGPPLCTYTLFLFCQVASFCCGYLSLWADDPVVDGQKMSALRGSIFGLGCLQAIG 1000

QY 1030 VFGYSMAVSIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFSEKELDVDSWIPE 1089

Db 1001 LPASMAAVFLGGARASCLLFRSLMDVARSPIGFFERTPVGNLLNRFSEKELDIVDVIDP 1060

QY 1090 VIKKMGSLFNVLGACIVILLATPIAIIPIPLGLIYFFVQRFYVASSRQLKRLSVRS 1149

Db 1061 KMTLLTYAFLGLEVLGSLVSMATPLAIVAILPLMLLYAGFQSLYVATCCQLRLESASYS 1120

QY 1150 PVYSHENETLLGVSVIRAFEEQRETHQSDLKVDENQKAYYPSIVANRWLAVLECVGNC 1209

Db 1121 SVCSHLAETFGSQVVRAPQAQGFPTAQHDALMDENQRISFPRLVADRVLAALELLGNG 1180

QY 1210 IVLFAALFAVISRHSLSAGLVLSVSYSLQVTTYLNLVMSSEMETNIVAVERLKEYSE 1269

Db 1181 LVFVAATCAVLSKAHLSAGLGSVSAAQLQVTTQLQWVRSWTDLENSMVAVERVQDYVH 1240

QY 1270 TEKEAPWQIQTAPPSPVQVGRVFRNTCLYREDLDLFLVLRHINVTINGEKVIGVRT 1329

Db 1241 TPKEAPWRLPSSAAQPLWPCGGQIEFRDFGLRHRPELPMVAVQVSLKIHAQKVGIVGRT 1300

QY 1330 GAGKSLTLGLPRINESABGEIIGDINIAKIGLHDLRFKTIIPQDPVLFSGSLRMNLD 1389

Db 1301 GAGKSLTWGLLRQLQATGGTWDGVPITDGLHTRLSRIIIPQDPVLFPSGLRMNLD 1360

QY 1390 PFSQVSDENVWTSLELAHLKDFVSLPDKLDHECAGBGENLSVGQRLVCLARALLRKT 1449

Db 1361 LLQENTDEGIWALETVOLKAFVTSPLGQLOVECSQGDLSVGQKOLLCLARALLRKTQ 1420

QY 1450 IIVLDEATAVDLETDLQSTIRTPQEDCTVLTAAHRLNTMDYTRVILDKBIOEYV 1509

Db 1421 IIVLDEATASVDPTGEIOMQALERWFAOCTVLLIAHRLRSYMNRCARVLVMDQGVAESG 1480

QY 1510 AFSDDLQORGLFYSMAKAGL 1530

Db 1481 SPAQLLAQKGLFYRLAQESGL 1501

RESULT 8

S51863

N:Alternate names: protein YCF1 - yeast (Saccharomyces cerevisiae)

C:Species: Saccharomyces cerevisiae

C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004

C:Accession: S51863; A55352; S50233

R:Oliver, K.; Harris, D.

A:Submitted to the EMBL Data Library, February 1995

A:Reference number: S51863

A:Accession: S51863

A:Molecule type: DNA

A:Residues: 1-1515 <OI>

A:Cross-references: UNIPROT:P39109; EMBL:248179; NID:g665657; PIDN:CAA88217.1; PID:g6656

R:Szyzypka, M.S.; Wemmie, J.A.; Moye-Rowley, W.S.; Thiele, D.J.

J. Biol. Chem. 269, 22853-22857, 1994

A:Title: A yeast metal resistance protein similar to human cyctic fibrosis transmembrane

A:Reference number: A55352; MUID:94357936; PMID:7521334

A:Accession: A55352

A>Status: nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-679, 'R', '681-1515 <SZ>

A:Cross-references: GB:L35237; NID:g556464; PIDN:AAA50353.1; PID:g556465

C:Genetics:

A:Gene: SGD:YCF1; MIPS:YDR135C

A:Cross-references: SGD:S0002542; MIPS:YDR135C

A:Map position: 4R

C:Function:

A:Description: required for cadmium resistance

C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology

C:Keywords: ATP; nucleotide binding; P-loop; transmembrane protein; yeast vacuole

F:287-308/Domain: transmembrane #status predicted <TM1>

F:345-366/Domain: transmembrane #status predicted <TM2>

F:421-442/Domain: transmembrane #status predicted <TM3>

F:446-467/Domain: transmembrane #status predicted <TM4>

F:534-555/Domain: transmembrane #status predicted <TM5>

F:558-580/Domain: transmembrane #status predicted <TM6>

F:546-829/Domain: ATP-binding cassette homology <ABC1>

F:663-670/Region: nucleotide-binding motif A (P-loop)

F:951-972/Domain: transmembrane #status predicted <TM7>

F:995-1016/Domain: transmembrane #status predicted <TM8>

F:1068-1088/Domain: transmembrane #status predicted <TM9>

F:1092-1113/Domain: transmembrane #status predicted <TM10>

F:1179-1200/Domain: transmembrane #status predicted <TM11>

F:1208-1229/Domain: transmembrane #status predicted <TM12>

F:1289-1483/Domain: ATP-binding cassette homology <ABC2>

F:1306-1313/Region: nucleotide-binding motif A (P-loop)

Query Match 28.3%; Score 2838; DB 1; Length 1515;

Best Local Similarity 40.3%; Pred. No. 2.8e-182;

Matches 632; Conservative 295; Mismatches 495; Indels 146; Gaps 27;

QY 28 DFTKCFQNTLVVWPCFYLMACFPFYF-----LYLSRHD-----RGVIOMTPLNKTXT 75

Db 27 DFTQCFIDGVIL-----NLSAIFMTTFGRLVNLCKKSGIKYRRNWIIIVSRVALVLL 81

QY 76 ALGFLWLVIVCWADLPYSFWERSRGIFLAPVFLVSTLIGTTLTFLTLQIERRKGVSS 135

Db 82 EIAF-----VSLASLNISKEAEN---FTIVSQYASTMLSLFVALAHWIEYDR--SVVAN 132

QY 136 GIMLTFWLVVALCALAILRSKIMTAKEDAQVDLFRDITFVYVES-----LLLIQLVLS 190

Db 133 TVLFLYFWLFEFGNPAKILNIRHYEG-----IWSGGTGFILTLFQVITC 180

QY 191 FS-----DRSPL-----FSETHDPNCPCESSASFLSRITFWMITGLIVRGYRQPLEG 238

Db 181 ASILLLEALPKPLMPHQHITLRRKPNVDSANIFSRITFSWMSGLMKTGYEKYLVE 240

QY 239 SDLASLNKEDTSEQVVPVLVGNWKKCAKTRKQPVVYSSKDPAPQKSSKVDANEVE 298

Db 241 ADLYKLPRNFSSSELSQLEKNWENL-----KQKS----- 271

QY 299 ALIVKSPQKWNPSLKVLYKTGPFYFLMSPFFFKAIHDLMMFSGPQILKLIKPVND--- 355

Db 272 -----NPSLSWAICETFGSKMLAAFAFKAIHDVLAFTQPOLRLIKFTVDYNS 320

QY 356 -----TKAPDMQGYFVTVLLFVTAQTLVLHQYFHCIFVSGMIRIKTA 398

Db 321 ERQDDHSSLOGFENNHPQKLPVIRGFLIAFAMFLVGFQTSLVHQYFLNVENTGMYIKSA 380

QY 399 VIGAVYRKALVITNSARKSSTVGEIVNLSVDAQRFMDLATYINMINSPLOVILALYL 458

Db 381 LTALIYOKSLVLSNEASGLSSTGDI VNLMSVDVQKLODLTQWLNLSWGPFOIICLYSL 440

QY 459 WNLGSPSVLAGVAVVNLMPVNVAMMKTQTYVAHMKSKDNRIKLNELINGIKVLKLY 518

Db 441 YKLLGNSMWGVIIIVIMPLNSFLMRKQKLOKSKYKQKDERTRVISEILNLIKSLKLY 500

QY 519 AWELAFDKVLAIR-QBELKVLKKSVALSVAGTTTWTCTPFLVALCTFAVVTVDENNIL 577

Db 501 AWEXFYREKLEFEVRNNKELNLTLCGYMAVTSFQFNVLPFLVSCCTFAFVY- YTEDRAL 559

QY 578 DAQTAFFVSLALFNILFRPLNLPWISSIVOASVSLAKRLRIFLSHELEPDSIRRRPVKD 637

Db 560 TTDLVFPALTFLNLSLFPMLIIPVNLSPFASVIGRLFTFTNEELQPDSDVQRLPKVK 619

QY 638 GGGTNSITV-RNATFTTWARSD--PPTLNGITFSPGALVAVVGQVCGKSSLSALLAE 694

Db 620 NIGDVAINIGDDATFLWKPVEYKVALKNINFOAKGNLCTIVGKVGSGKTALLSCMLGD 679

QY 695 MDKVEGHVAIKGSVAVYVPOQAMQNDLSRENILFGCQLEBPYRSVTSVQACALLPDLLEIP 754

Db 680 LFRVKGFPATVHGSVAVYVQVFWMNGTVKENILFGHRYDAEFYEKTIKACALITDLAILM 739





QY 1227 AGLVGLSVSYSLQVTTYNLWLVMSSEMETNIVAVERLKEYSEKEAPMOIQETAPSS 1286  
 DB 1173 SGLVGLSLSYAVQITQSLTFVVRQSDVETNIVSVERMLEYIGLPSEAPSIIPOHRPPEG 1232  
 QY 1287 PVOGRVFRNYCLRYREDLOPVLRHINVTINGGKVGIVGRTGAGKSSLTGLGFRINES 1346  
 DB 1233 WPSHGAIFDHVSRYRENPLVNDISVNIKQEKIGIVGRTGAGKSTLTALFRLIEP 1292  
 QY 1347 ABETIIIDGINTAKTGLHDLRPKITIIPQDPVLFSGSLRPMNLDPDSQVSDREVWTSLELA 1406  
 DB 1293 TSGDQLDDINITSIGLHDLRSRLAIIPQENQAFEGTIRENLDPNANATDEEIHAALEAA 1352  
 QY 1407 HLKDFVSALPDKLDECEGEGNSVGORQLVCLARALLRTKILVLDEATAAVDLETD 1466  
 DB 1353 SUKQFIQTLDGGLYRVTGEGANLSSGQQLMCLTRALLTPTRVLLDEATAVUVEDA 1412  
 QY 1467 LTQSTIRTFEDCTVLTIAHRLNTIMDYTRVTLVDKGEIQEYAGPSDILQOR-GLFYDMA 1525  
 DB 1413 IVQRTIRERFNDRTILTIAHRLNTIMVDSNRILVDHGKVFEDSTKLEENKASLFYSLA 1472  
 QY 1526 KDAGLV 1531  
 DB 1473 KESGLI 1478

RESULT 11  
 T01369  
 ABC transporter AtMRP2 [imported] - Arabidopsis thaliana  
 N:Alternate names: multidrug-resistance protein homolog T29F13.13  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C>Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 09-Jul-2004  
 C:Accession: T01369; D84759  
 R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul  
 submitted to the EMBL Data Library, May 1998  
 A:Description: Arabidopsis thaliana chromosome II BAC T29F13 genomic sequence.  
 A:Reference number: Z14179  
 A:Accession: T01369  
 A>Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-1623 <CROU>  
 A:Cross-references: UNIPROT:O64590; EMBL:AC003096; NID:g3132469; PID:g3132479  
 A:Experimental source: cultivar Columbia  
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;  
 euse, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
 Nature 402, 761-768, 1999  
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
 A:Reference number: A84420; MUID:20083487; PMID:10617197  
 A:Accession: D84759  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-1623 <STO>  
 A:Cross-references: GB:AE002093; NID:g3132479; PIDN:AAC16268.1; GSPDB:GN00139  
 C:Genetics:  
 A:Gene: At2g34660; T29F13.13  
 A:Map position: 2  
 A:Introns: 110/3; 170/2; 183/3; 236/1; 274/2; 301/2; 309/3; 328/3; 350/3; 375/3; 411/3;  
 C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology  
 F;1252-1452/Domain: ATP-binding cassette homology <ABC>

Query Match 23.8%; Score 2384.5; DB 2; Length 1623;  
 Best Local Similarity 35.2%; Pred. No. 1.1e-151;  
 Matches 561; Conservative 274; Mismatches 534; Indels 223; Gaps 30;  
 44 FYWLACPPF-YFLVLSHRDGYQMPPLNTKTLG---FLMIVCWADLFYSFW----- 94  
 5 FIEWYKFPVNGVMTQVANAFGAYTPCATDSFVLGISQLVLLVLC-----LYRIWLAKD 60  
 95 -----ERSRGIFLA-----PVFLVSPTLIGITL-----LATFLIQLERRK 130  
 61 HKVERFCRLRLNYFLALLAAVATAEPLRL-----INGISVLDGFGCLPPF----- 109  
 131 GVOSGIMLTFWLVALVICALAILRSKIMTALKEDAQVDLFRDITFYVYFS-----LL 182

DB 110 --EAFGLGVKAFANGAVMMILMETKI-----YIRELRWYVRFVAVIALVGDV 156  
 QY 183 LIQVLVS-----CFSDRSPLFSETHD--- 204  
 DB 157 LLNLVLVSKEYSYVLYLYTSEVGAQVLFGLFMHLPNDLTPGYNPVRSETHD--- 216  
 QY 205 -----ENPCESSASPLSRITFWITGLVIRGVRQPLEGSDLSLNKEDTSEQVVPVLV 258  
 DB 217 EETSDGOQIICPEKHANIFDKIFFSMWNPMTLGSKRPLTEKDVWYLDTDQTELTFTSFQ 276  
 QY 259 KXWKECAKTRKQPVVYSSKQDPAQKESKVDANEVEALIVKSQKWNPSLFKVLV 318  
 DB 277 HSWDKELQK-----PQ-----PWLRLALN 295  
 QY 319 KTFGPYFLMSPFFKAIHDLMMFSGPQILKLLIKFVNDTKAPDWOGYFVTVLLFTYACLOT 378  
 DB 296 NSLGGRFWVGFWKIGNDSCQFVGPLLLNLQKMSQB-DAPAMWGYIYAFSIFVGVVFGV 354  
 QY 379 LVLHQYFHCIVSGMRIRKTAIVIGAVYRKALVITNSARKSTVGEIVNLMSVDAQRFMDLA 438  
 DB 355 LCEAQYFQNVNRVGYRLRSALIAAVERKSLRLTNEGRKFTQKITNLMTTDAESLQOIC 414  
 QY 439 TYINWISAPLOVILALYLLMLNLPSPVLGAVAVMLVFNVAVMAMTKTYQVAHMKSK 498  
 DB 415 QSLHTMWSAPFRIIITILYQQLGVASLIGALLLVLMFPPLQTVIISKQKLTKEGQRT 474  
 QY 499 DNRIKLMNEILNGIKVLKYAWELAFKDKVLAIRQELKVLKKSAYLSAVCTFTWCVTFP 558  
 DB 475 DKRIGLMNEVLAAMDVTVCAYAWNSFQSKVTVTRDDLSWFRKSKQLLALNMFLNSIPV 534  
 QY 559 LVALCTFAVYVITDENNILDQAFAVSLALFNILRFLNLPIMVSISSVQASVSKRL-- 616  
 DB 535 LVTIVSGVFTLLGD--LTPARAFTSLSLFAVLRFLPLFPLNIIITQVNVANVSLKREE 592  
 QY 617 -----RIFLSHEELEPDSIERPVRKVGCGGTNSITVRNATFTW-ARSDPTLNGITFTSIP 669  
 DB 593 VLATEERILPNPIEP-----GPASIRNGYFSWDSKGRDPTLSNLINLNDVP 640  
 QY 670 EGALVAVVGVQVCGKSSLSALLAEMDKV-BGHVAIKGSVAVVPOQAQWIONDSLENI 728  
 DB 641 LGLSVAVVGVSTGEGKTSLSAILGELPATSDAIVTLRGSVAVVPOQVSIENATVRDNL 700  
 QY 729 GQCLEEPPYRSVIOACALLPLEILPSGDRTEIGKGNLSCGQKQVSLARAVYSNADI 788  
 DB 701 GSPDREKYEIRADIVTSLKHLLELLPLPGDLTEIGERYNIGSGQKQVSMARAVYNSDV 760  
 QY 789 YLFDPLSAVDHGVCKHTFENVIGPKMLKNKTRITLVTHSMSYLPQVDVIVMSGKISE 848  
 DB 761 YIFDDPLSALDAHVGQVQVFEKI--KRELQKTRVLVTNQLHFLSQVDRIVLVHSGTVKE 818  
 QY 849 MGSYQELLARDCAFAEFLRTYASTEQEQDAENGVTGSGPKE--AKQMENGMLVTDQA 906  
 DB 819 EGTVEELSSNGPLFORLMEAGKVEEY--SEENGAEADQTAEQAPVANGNTNGLQMGSD 876  
 QY 907 GQKQORQLSSSSSYSGDISRHNSHTAELOKAEKETEKLMEADKAQTKGVKLSVYWDY 966  
 DB 877 DKK-----SKENKKGKGSVLKIKQBERETGVVSWRVLRKY 911  
 QY 967 MKAI-GLFISFLSIFLEMCNHYVALASNVLSLWTDTPVNGT-QEHYTKVRLS-VYGALG 1023  
 DB 912 QDALGAGVMMMLLCYVLTEVFRVTSSTWLSWTD-----AGTPKSHGPLFYNLYALLS 967  
 QY 1024 ISQGIAGVFGYSNAVSGIGTILASRCLHLDLHLSILRSPMSFFERTPSGNLVNRPSELOTV 1083  
 DB 968 FQGVLTLTNSYWLIMSSLYAAKLDHNLHLSILRAPMSFFHTNPLGRINRFAKDLGI 1027  
 QY 1084 DMSIEVIMKMGSLFNVTGACIVILLATPIAAIIPPLGLIYFFQVREYVASSRLKRL 1143  
 DB 1028 DRTAVAVNMFMFGVQSLLSTVYLGIVSTLSLWAIMPLLVLFYGAIFYQNTAREVKRM 1087  
 QY 1144 ESVRSRPFVSHNETLLGVSVIRAFEEQERFTHQSDLKVDENQKAYPYSIVANRWLAVRL 1203



Db 1232 PSSGSIKPEDVVLRYRPELPPVLLHGVSLISPMKVGIVGRTGAGKSSLLNALFRIVELE 1291  
 Qy 1348 EGEIIDDGINAKIGHDLRFKTIIPQDPVLPFGSLRNWLDPPFSYDEEYVWTSLELAH 1407  
 Db 1292 KRIILDECDIGRGLMDLRKVLGIIPQAPVLPFGVFRNLDPPFSHNDQDLWESLERAH 1351  
 Qy 1408 LKDFVSALPKDHECAEGENLSVGORQLVCLARALLRKTILVLDEATAAVDETDDL 1467  
 Db 1352 LKDTIRNPLGLDAEVEAGENFSVGORQLLSARALLRRSKILVLDEATAAVDVRTVL 1411  
 Qy 1468 IQSTIRTOEDCTVLTIAHRLNTIMDYTRIVLVDKGEIQYGA PSDLLQORGLFYMSAXD 1527  
 Db 1412 IOKTIREEPKSCMLTIAHRLNTIIDCDKVLVDGKVGQSPSPENLLSNGESSFSK--- 1468  
 Qy 1528 AGLVGGGGGMLSRKGIIPPEYVLTRLAEDPAEPRYRTRE 1567  
 Db 1469 -----MQVSTGANAAYLRSITLENK-----RTR 1493

RESULT 13  
 E86428  
 probable ABC transporter [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
 C:Accession: E86428  
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
 anen, N.P.; Hughes, B.; Huizar, L.  
 Nature 408, 816-820, 2000  
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,  
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A:Reference number: A86141; MUID:21016719; PMID:11130712  
 A:Accession: E86428  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-1495 <STO>  
 A:Cross-references: UNIPROT:Q9C8H0; GB:AE005172; NID:g11055818; PIDN:AG28288.1; GSPDB:G  
 C:Genetics:  
 A:Map position: 1  
 C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology

Query Match 23.2%; Score 2326; DB 2; Length 1495;  
 Best Local Similarity 34.3%; Pred. No. 8.6e-148;  
 Matches 541; Conservative 314; Mismatches 566; Indels 154; Gaps 28;

Qy 2 ALRGFCS--ADGSDPLMDNNTWNTSNPDFTKCFQNTVLVWVP-----CFY-LWACPP 51  
 Db 5 ALNWKCPVADG-----WEKAVDGAFGAYTCADSLVWLVSHFVLGLCPYRIWIP- 58  
 Qy 52 FYFLYLSRHRDGYQMTPLNKTALGPLLWICVAD-----LPYFWSRGRGIFLAPV 105  
 Db 59 -----HNTKAIQVYLRKKYCNVLG-LACVCVPEPVLRVLMGSLFDMDDEDFDPF 110  
 Qy 106 FLVSPILLGITLLATFLQLERRKQVQSSGIMLFWLVVALVCAILRSKIMTALKEDA 165  
 Db 111 EVASLMVEAFAMFSLMVLGLGTEKQYKFRVYRFGVLVILVADAVLLDLVLPLKNSIN 170  
 Qy 166 QVDLFDIT---FYVYFSLLLQLVLCFSDRSPLFSETIHPD-----NPGC 209  
 Db 171 RTALVLFISRCQALFGILLIYIPEL--DYPGCHIYVNEPLDNVEVDALRGEGHICP 228  
 Qy 210 ESSASPLSRIITFWITGLIVRGQPLEGDLWSLNKEDTSQGVVPLVKNWKKCAKTR 269  
 Db 229 ERHASIFSRIFYGWIITPLMOLGYRKPITEKDVQWLDKWDQETLIKRFCORCWTESSRR-- 286  
 Qy 270 KQPVKVVYSSKQPAQPKSSKVDANEEVAILVSKPQKWNLSFKVLYKTGPFVLMGF 329  
 Db 287 -----PK-----FWLLRALNNSLGRFWLAG 307

Qy 330 FFKAIHDLMMFSSPOILLKLLIKFVNDTKAPDMQGYFYTVLLFVFTACLOTLVLRHQPHICF 389  
 Db 308 IFKIGNDLSQFVGPVILSHLLSRMQEGD-PAWGVYVYAFIIFVGVTLGLVCEAQIFQNVW 366  
 Qy 390 VSGMIKTAIVAGVYRKALVITNSARKSTVGEIVNLMSVDAQRFMDLATYTNMNIWSAPL 449  
 Db 367 RVGPFRLSTLVAAILFKSLRLTHEARKNFASPKVTNMTITDANALQQIISQOLHGLWSAPP 426  
 Qy 450 QVILALYLMLNLGFSVLGAVAVMLVMPVNAVMAKTKTYQVAHMKSKDNRIKLMNEIL 509  
 Db 427 RIIVSMILLYQQLGVASLFGSLILFELIPLQTLIISKRKLTKEGLQWTDKRVGITEIL 486  
 Qy 510 NGIKVILKYAWELAFKDKYLAIHQBELKVLKKSAYLSAVGTFTWCTPFLVALCTFVYV 569  
 Db 487 SSMDTVKCYAWEKSPESRIGQIRNEELSWFRKAQLLSAFNSFILNSIPVVTWVSGVFP 546  
 Qy 570 TIDENNILDAQTA FVSLALFNILRPILNLPWISSIVOASVLSKRL-RIFLSHELEPDD 628  
 Db 547 LLGGD--LTPARAFSTLSLFAVLRPPLNMLPNLLSQVNVNANVSLQRIEELLSEERILAQ 604  
 Qy 629 STERRPVKGGGTNSITVRNATFTW-ARSDPPTLNGITPSIPEGALVA VVGQVGGCKSSL 687  
 Db 605 NPPLQD-----GTPAISIXNGYFSDSKTKTPTLSDINLEIPVGTVAIVGTGEGKTSL 659  
 Qy 688 LSALAEAMDKVE-GHVAIKGSVAYVPOQAQIOWNSLRENILFGCQLEEPYRSVIOACAL 746  
 Db 660 ISAMIGELSHAETTSVWIRGVSVAVYVQVSWIFNATVRENILFGSDFESERYWRAIDATAL 719  
 Qy 747 LPDLBILPSGDRTEIGEGVNLGSGQKORVSLARAVYSNADIYLFDDPLSAVDHVGKHI 806  
 Db 720 QHDLDFGRDLTEIGERGVNISGQKORVSNARAVYSNDSYIYIFDDPLSALDAVHAQV 779  
 Qy 807 FENVIGPQKMLKNKTRILVTHSMSVLPQVDVIVMSGGKISBMSVSOELLARDGAFABFL 866  
 Db 780 FDSCH--KDELRGKTRVLVTNQLHFLPLMDKILVSEGMIKEGTFVELSKSGILFKKUM 837  
 Qy 867 RTYASTEQEQAEENGVTGVSQPGKEAKQEMENGLVTDTSAGKQLQQLSSSSSYSGDISR 926  
 Db 838 ENAGKMDATQEVNTN-----DENILKLGPTVTVDVS---ERNLGSTK----- 876  
 Qy 927 HINSTAELOKABAKKEETKLMLEADKAQGTQVKLSVYNDYMKAI-GLFTSIFSLIFLPMCN 985  
 Db 877 -----QGKERRS-VLIQKEERETGIISWVLMRYKEAVGGLVWVMIILLACYLAT 924  
 Qy 986 HVSALASNVWLSLWTDPIVNGTQEHK-----VRLSVYVYALGIGSQIAVFGVSMASVIG 1040  
 Db 925 EVLRVSSSTWLSIWTD-----QTSKNYSPGFYIVVYALLGFGQVAVTFNNSFWLITS 977  
 Qy 1041 GILASRCLHVDLLHLSILRSPMSFFERTPSGNLVNRPFSKELDTVDSMIPVIMPFMGSLFN 1100  
 Db 978 SILHAARLHDAMLSSILRAPMLFFHTNPTGRVINFSKIDIGDIRVANLMMNMFNQLAQ 1037  
 Qy 1101 VIGACIVILLATPIAAIIPPLGLIYFFVQRPYVASSRQLKLESVSRSPVTSFHNFTULL 1160  
 Db 1038 LLSTPALICTVSTISLWAIMPLLIIFAYAYLYQSTRSREVRDLSDTSRSPYIAQGEALN 1097  
 Qy 1161 GVSIVRAFEEOERFIHQSDLVKVDENQKAYPSIVANRMLAVRLECVGCNCIVLPAALFAVI 1220  
 Db 1098 GLSSIRAYKADYRMAKINGKSMNNIRFTLANTSSNRMLTIRLETIGGVMIWLTATFVIL 1157  
 Qy 1221 SRHSLS-----AGLVGLSVYSLOVTTYLNMLVRMSSEMETNIVAVERKEYSETEKEAP 1275  
 Db 1158 QNGNTNNAQFASWTGLLSYLTNITSLLSGLVLRQASRAENSLNSVERVGNVLDLPSEAT 1217  
 Qy 1276 WOIQETAPPSNWPQVRVEFRNYCLURYREDLDFVLRHINVTINGEKVIGVORTGAGKSS 1335  
 Db 1218 DIENNRVPCWPGSGSIFKFDVHLRPLPPLVHLGLTFFVSPSEKVGWVGTGAGKSS 1277  
 Qy 1336 LTLGIFRNESAGEIIDIIGNIAKIGHDLRFKTIIPQDPVLPFGSLRNWLDPPFSYDEEYVWTSLELAH 1395  
 Db 1278 MLNALFRIVEKEGRIMIDDCDVAKFGTUDVRVLISIIIPQSPVLFSGTFRFNIDPFSEHN 1337



A:Residues: 1-255 <BOY>  
A>Note: the authors translated the codon CAG for residue 248 as His  
R:Purnelle, B.; Goffeau, A.  
submitted to the EMBL Data Library, April 1996  
A:Description: The sequence of 32 kb on the left arm of yeast chromosome XII reveals 14  
family and a new ABC transporter homologous to the human multidrug resistance protein.  
A:Reference number: S69380  
A:Accession: S69391  
A:Molecule type: DNA  
A:Residues: 1-1559 <PUR>  
A:Cross-references: EMBL:X97560; NID:g1297003; PIDN:CAA66162.1; PID:g1297015  
R:Miosga, T.; Zimmermann, P.K.  
Yeast 12, 693-708, 1996  
A:Title: Sequence analysis of the CEN12 region of Saccharomyces cerevisiae on a 43.7 kb  
e conductance regulator protein CFTR.  
A:Reference number: S70557; MUID:96405918; PMID:8810043  
A:Accession: S70560  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-1559 <MIW>  
A:Cross-references: EMBL:X91488; NID:g1495203; PIDN:CAA62776.1; PID:g1495208  
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1995  
C:Genetics:  
A:Gene: SGD:BPT1; MIPS:YLI015W  
A:Cross-references: SGD:S0003938  
A:Map position: 12L  
C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology  
C:Keywords: ATP; nucleotide binding; P-loop; transmembrane protein  
F:28-44/Domain: transmembrane #status predicted <TM1>  
F:143-159/Domain: transmembrane #status predicted <TM2>  
F:178-194/Domain: transmembrane #status predicted <TM3>  
F:334-350/Domain: transmembrane #status predicted <TM4>  
F:421-437/Domain: transmembrane #status predicted <TM5>  
F:526-542/Domain: transmembrane #status predicted <TM6>  
F:550-566/Domain: transmembrane #status predicted <TM7>  
F:654-847/Domain: transmembrane #status predicted <TM8>  
F:672-679/Region: ATP-binding cassette motif A (P-loop)  
F:974-990/Domain: transmembrane #status predicted <TM9>  
F:1017-1033/Domain: transmembrane #status predicted <TM9>  
F:1099-1115/Domain: transmembrane #status predicted <TM10>  
F:1118-1134/Domain: transmembrane #status predicted <TM11>  
F:1212-1228/Domain: transmembrane #status predicted <TM12>  
F:1319-1529/Domain: ATP-binding cassette homology <ABC2>  
F:1336-1343/Region: nucleotide-binding motif A (P-loop)

Query Match 22.9%; Score 2294.5; DB 1; Length 1559;  
Best Local Similarity 35.1%; Pred. No. 1.2e-145;  
Matches 574; Conservative 292; Mismatches 541; Indels 227; Gaps 39;

QY 23 NTSNPDFTKCFQNTVLVWVPCFYLW-ACPPFYFLYLSRHRDGYIOMTPLNKTALGFL- 80  
DB 23 NALNP-----CFISVISAWQAVFLLIGSYQLMKLYKNNKVPFRFKNFPLPDKINSRHLT 78

QY 81 -LMIVCW-----ADLFYSFWSRSGIFLAPFVLVSPITLGI---TTLAT 121  
DB 79 HLTVNCFQSTLIICELALVQSDDRVYFPLK-KALYLNLLF-----NLGISLPTOYLAY 132

QY 122 FLIQLERKGVQSGGIMLTFWLVALVCALAILRSKI-----MTALKEDAQV 167  
DB 133 F-----KSTFSGMGNLFYMFQILLQLFLILQRYHSGSNRLTVISQGTAMILEVLL 185

QY 168 DLFRDITFYVYFSLILLQLVLSCFSDRSPLFSTIHDNPNCPSSASFLSRITFWMITGL 227  
DB 186 -LNSVAIFIY-DLCIFEPI-----NELSEYKKNQWY-----PVHVLSYITFIWKNL 233

QY 228 IVRGYRQPLEGSDLWSLUNK---EDTSEQVVPVLVKNWKECAKTRKOPKVVVSSKDAQP 285  
DB 234 IVETIR-----NKKIKDNQLPLPVDLNIK----- 259

QY 286 KESSKVDAHEVEALIVKQKWNPSLKVLYKTFGPFLMFFFKAIHDLMMFSGPQI 345  
DB 260 SISKEFKANWELEKWLNRN-----SLWRAIWKSGFRTTISVAMLYETTSDLLSVVQFQ 312

QY 346 LKLLIKFVN---DTKAPDMQGYFYTVLLFVTAQLTQVLVHQYHFHCFVSGRIKTAVIGA 402  
DB 313 LRIFDGLNPETSKYPPPLNGVFIALTLFVISVVSFLTNQFYIGIEAGLIRGSLASL 372

QY 403 VYRKALVITNSAKSGSTVGIVNLSVDA---ORPMDLATYINWINSAPLOVALALYLLW 459  
DB 373 VYQKSLRLTLAERNEKSTGDLNLSVDVLIQIRFFENA---QTIIGAPIQIIVLVTSLY 429

QY 460 LNLGPSVLGAVVAVVLMVVPVNAVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKLYA 519  
DB 430 WLKGVAGIGLVLTWALMMPINAFLSRKVKKLSKTKMYKDKRIKTIITELLNAIKSLYA 489

QY 520 WELAFKDKVLAIROE-ELKVLKKSAYLSAVGTFTWCTPFLVALCTFAVVTIDENNILD 578  
DB 490 WEEPMARLNRHVRNDEMELNFKRIGIVSNLIYFAWNCVPLMVTGCTGLF-SLFSDSL 548

QY 579 AQTAFLVALFNILRPFPLNLPWVSISSVQASVLSKELRIFLSHEEPEPSIER-RPVK 637  
DB 549 PAIVFPSLSLNFILNSAIYSVPSMINTIITSVSMERLKFLLSDEIDDSFIEDPSAD 608

QY 638 GGGTNSITVRNATFTW-----ARSDPPTLNGIT-----FSIPEGALV 674  
DB 609 ERALPALTEMNITFLMKSKEVLTSQSGDNLRTDEESIGSSQIALKNIDHFEAKRGDLV 668

QY 675 AVGVQVCGKSSLLSALLAEMDKVEGH-----VAIKGSVAVVPOQAMIQNDLSRENI 726  
DB 669 CVVGRVGAGKSTFLKAILGQLPCMSSGRDSIPPKLIIRSSSVAYCSQESWIMNASVRENI 728

QY 727 LFGQLEBPYRYSVIQACALLPDLIELPSGDRTEIGKGVNLSGGOKORVSLARAVYNA 786  
DB 729 LFGHKFDQYDYDLTIKACQLLPDLKILPDGDETUVGKGLSLSGGQKARLSARAVYRA 788

QY 787 DYLFDPLSAVDVAHVGHKIFENV-IGPKGMLKNKTRILVTHMSYLPQVDVIVMSGGK 845  
DB 789 DYLDDILSAVDAEVSNIIEYVLIGTKALLKNKTIILNTVTSILKHSQMIYALENGE 848

QY 846 ISEMGVSOELLAR---DGAFAEFLRTVAS-----TEQODAEENGVTGVSFGKEAKQME 897  
DB 849 IVEQGNVEDVMNRKNNTSKLKLLEEFDSPIDNGNESDVQTEHRSSEVDEP----- 900

QY 898 NGMLVDSAGKQLQRLSSSSSYSGDISRHNSHNSAEQLQKABAKKEETWKL----- 947  
DB 901 -----LQKVTSETEDEVVT-----ESELIKANSRRASLATLAPRPFVGAQL 945

QY 948 -----MEADKAQTQGVKLSYMDYMKALGLFISFLSIFLPM-CNHVSALASNYWLSLW 999  
DB 946 DSVKTAQKAETEVGRVKTILYLAIKCGV-LGVWLFLLFMLTRVDFDLAENFWLKYW 1004

QY 1000 TDDPIVNGTOEHTKVRVLYGALGISQGIAPVGYSMAYSI---GGILASRCLHVDLLHSI 1056  
DB 1005 SESNEKNGSNRVMFVGVYSLIGVAS--AAFNNLRSIMMLLYCSIRGSKKLHESMAKSV 1062

QY 1057 LRSNMSFERPSPGNLNRFSKELDTVDSPITPEVIKMFMGSLFNIVGACIVILLATPIAA 1116  
DB 1063 IRSPMTFTETTPVGRHIIIRFSDMDVDSNLQYIFSFYFKSILTLVTLVTVILVGNWPL 1122

QY 1117 IIPPLGLIYFFVQRFYVASSRQLRLESVRSRSPVYSHNFNTLLGVSVIRAPEROERFIH 1176  
DB 1123 VENMELVVIYIYQTFYVLSRELKRLISISYSPIMSLMSESLNGYSIIDAHDHERFIY 1182

QY 1177 QDQLKVDENQKAYYSPISIVANRWLRLSCVGNICVLPAALPAVI---SRHSISAGLVGLS 1233  
DB 1183 LNYEKIQVNVDFVFNFRNTRNLRSVRLQITIGATILATAILATATWNTKQLSSGWGLL 1242

QY 1234 VSYSLQVTTYLNWLRMSSEMETNIVAVERLKEYSETKEAPWQIQETAPSSSWPQVGRV 1293  
DB 1243 MYSLEVTSGLTWIVRTVTIETNIVSERIVEYCELPPEAQOSINPEKRPDENWPSKGI 1302

QY 1294 EFRNYCLRYREDLDFVLRHINVTNGEKGIVIGRTGACKSLITLGLFRINSEAGEIIL 1353  
DB 1303 EFKNYSTKYRENLDPLVNLNNVKBPCKEKVGIVGRTGACKSTLSLALFLEPTEGKIIL 1362

QY 1354 DGINIAKIGLHDLRFKTIIPQDPVLFSGSLRNWLDPPFSQYSDSEEVWTSLELAHLKDFVS 1413

Db 1363 DGIDISDIGLFLRSHLAIIPQAFEGTVKTNLDPFNRYSEDELKRAVEQAHLKPHLE 1422  
QY 1414 AL-----PDKLDEHCAEGENLSVGORQLVCLARALLRKTILVILDEA 1456  
Db 1423 KMLHSPRGDDSNEDGNVNDILDVKINENGSLVGORQLCLARALLNRSKILVILDEA 1482  
QY 1457 TAAVDLETDDLIQSTIRTOPEDCTVLTIAHRLNTIMDYTRVIVLDKGEIOEYCAPSDLIQ 1516  
Db 1483 TASVDMETDKIIQDTIRREFKORTILTIAHRIDTVLDSKIIIVLDQGSVREFDPSKLLS 1542  
QY 1517 QR-GLFYSMAXDAG 1529  
Db 1543 DKTSIFYSLCEKGG 1556

Search completed: March 18, 2005, 11:05:44  
Job time : 62.7349 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model  
Run on: March 18, 2005, 10:57:20 ; Search time 206.058 Seconds  
(without alignments)  
4838.524 Million cell updates/sec

Title: US-10-665-283-8  
Perfect score: 10016  
Sequence: 1 MALRGFCSADGSDPLDWNV.....RSVAVAKPKFISPDLS 1947

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt 03: \*  
1: uniprot\_sprot: \*  
2: uniprot\_trembl: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7860	78.5	1531	1 MRPI_HUMAN	P3527 homo sapien
2	7769	77.6	1515	2 Q9UQ99	Q9ug99 homo sapien
3	7695	76.8	1531	2 Q864R9	Q864r9 macaca fasc
4	7689	76.8	1531	2 Q864S0	Q864s0 macaca fasc
5	7444	74.3	1459	2 Q9UQ97	Q9ug97 homo sapien
6	7419.5	74.1	1456	2 Q9UQAO	Q9ugao homo sapien
7	7272	72.6	1531	2 Q6UR05	Q6ur05 canis famil
8	7230.5	72.2	1530	2 Q8HX05	Q8hx05 bos taurus
9	7104.5	70.9	1400	2 Q9UQ98	Q9ug98 homo sapien
10	7002.5	69.9	1528	1 MRPI_MOUSE	Q35379 mus musculu
11	6932.5	69.2	1532	2 Q810E4	Q810e4 rattus norv
12	6932.5	69.2	1532	2 Q8CG09	Q8cg09 rattus norv
13	6892	68.8	1523	2 Q810G9	Q810g9 rattus norv
14	6075	60.7	1215	2 Q68CP7	Q68cp7 homo sapien
15	4522	45.1	1519	2 Q80ZK8	Q80zk8 mus musculu
16	4491.5	44.8	1527	1 MRPI_HUMAN	O15438 homo sapien
17	4422	44.1	1514	2 Q96QA9	Q96qa9 homo sapien
18	4396	43.9	1522	1 MRP3_RAT	Q88563 rattus norv
19	3952.5	39.5	1592	2 Q7Q1D7	Q7q1d7 anopheles g
20	3885.5	38.8	1548	2 Q7KTC3	Q7ktc3 drosophila
21	3880.5	38.7	1548	2 Q7KTC1	Q7ktc1 drosophila
22	3879.5	38.7	1548	2 Q7KTB7	Q7ktb7 drosophila
23	3874.5	38.7	1548	2 Q7KTC2	Q7ktc2 drosophila
24	3860.5	38.5	1548	2 Q917N0	Q917n0 drosophila
25	3857	38.5	1549	2 Q7KTC0	Q7ktc0 drosophila
26	3845.5	38.4	1548	2 Q8T9C5	Q8t9c5 drosophila
27	3838.5	38.3	1548	2 Q7KTB8	Q7ktb8 drosophila
28	3836.5	38.3	1548	2 Q7KTB9	Q7ktb9 drosophila
29	3818.5	38.1	1548	2 Q7KTD0	Q7ktd0 drosophila
30	3813.5	38.1	1548	2 Q7KTC8	Q7ktc8 drosophila
31	3812.5	38.1	1548	2 Q7KTC4	Q7ktc4 drosophila

RESULT 1		ALIGNMENTS	
ID	MRPI_HUMAN	STANDARD;	PRT; 1531 AA.
AC	P3527; O14819; P78419;		
DT	01-FEB-1994 (Rel. 28, Created)		
DT	01-NOV-1997 (Rel. 35, Last sequence update)		
DT	25-OCT-2004 (Rel. 45, Last annotation update)		
DE	Multidrug resistance-associated protein 1 (ATP-binding cassette, sub-family C, member 1).		
GN	Name=ABCC1; Synonyms=MRP, MRPI;		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RP	MEDLINE=9308080; PubMed=1360704;		
RA	Cole S.P.C., Bhargwa J.G., Gerlach J.H., Mackie J.E., Grant C.E.,		
RA	Almquist K.C., Stewart A.J., Kurz E.U., Duncan A.M.V., Deeley R.G.;		
RT	"Overexpression of a transporter gene in a multidrug-resistant human lung cancer cell line."		
RT	Science 258:1650-1654(1992).		
RL	[2]		
RP	REVISIONS.		
RP	MEDLINE=93262415; PubMed=8098549;		
RA	Cole S.P.C., Deeley R.G.;		
RT	"Multidrug resistance-associated protein: sequence correction."		
RL	Science 260:879-879(1993).		
RP	[3]		
RP	SEQUENCE FROM N.A.		
RP	MEDLINE=96251691; PubMed=8649356;		
RA	Stride B.D., Valdimarsson G., Gerlach J.H., Wilson G.M., Cole S.P.C.,		
RA	Deeley R.G.;		
RT	"Structure and expression of the messenger RNA encoding the murine multidrug resistance protein, an ATP-binding cassette transporter."		
RT	Mol. Pharmacol. 49:962-971(1996).		
RP	[4]		
RP	SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.		
RP	MEDLINE=98008927; PubMed=9344662; DOI=10.1006/geno.1997.4950;		
RA	Grant C.E., Kurz E.U., Cole S.P.C., Deeley R.G.;		
RT	"Analysis of the intron-exon organization of the human multidrug-resistance protein gene (MRP) and alternative splicing of its mRNA."		
RT	Genomics 45:368-378(1997).		
RP	[5]		
RP	SEQUENCE OF 1131-1531 FROM N.A.		
RP	MEDLINE=99425270; PubMed=10493829; DOI=10.1006/geno.1999.5927;		
RA	Lofthus B.J., Kim U.-J., Sneddon V.P., Kalush F., Brandon R.,		
RA	Fuhrmann J., Mason T., Crosby M.L., Barnstead M., Cronin L.,		
RA	Deslattes Mays A., Cao Y., Xu R.X., Kang H.-L., Mitchell S.,		
RA	Eichler E.B., Harris P.C., Venter J.C., Adams M.D.;		
RT	"Genome duplications and other features in 12 Mb of DNA sequence from human chromosome 16p and 16q."		
RT	Genomics 60:295-308(1999).		
RL	[6]		

32	3807.5	38.0	1548	2	Q7KTC9
33	3793.5	37.9	1548	2	Q7KTC7
34	3790	37.8	1549	2	Q9VK56
35	3771.5	37.7	1548	2	Q7KTC5
36	3769.5	37.6	1548	2	Q7KTC6
37	3741	37.4	1564	2	Q8QG98
38	3738.5	37.3	1567	2	Q6PH26
39	3655	36.5	822	2	Q9JHS0
40	3653.5	36.5	1544	2	O6PSM3
41	3641.5	36.4	1544	2	Q9SM36
42	3625.5	36.2	796	2	Q8C7V6
43	3595	35.9	1545	1	MRP2_HUMAN
44	3594	35.9	1544	2	Q95L75
45	3577	35.7	1564	1	MRP2_RABIT

Q7ktc9 drosophila
Q7ktc7 drosophila
Q9vk56 drosophila
Q7ktc5 drosophila
Q7ktc6 drosophila
Q8qg98 raja erinac
Q6ph26 brachydanio
Q9jhs0 rattus norv
O6psm3 canis famil
Q9sm36 canis famil
Q8c7v6 mus musculu
Q92887 homo sapien
Q95l75 macaca mula
Q28689 o canalicul

TOPOLOGY, AND CARBOHYDRATE-LINKAGE SITES.  
 RX MEDLINE=97442425; PubMed=9295302; DOI=10.1074/jbc.272.38.23623;  
 RA Hipfner D.R., Almqvist K.C., Leslie E.M., Gerlach J.H., Grant C.E.,  
 RA Deeley R.G., Cole S.P.C.;  
 RT "Membrane topology of the multidrug resistance protein (MRP). A study  
 RT of glycosylation-site mutants reveals an extracytosolic NH2  
 terminus.";  
 RL J. Biol. Chem. 272:23623-23630(1997).  
 RN [7]  
 RP TOPOLOGY.  
 RX MEDLINE=97476249; PubMed=9334225; DOI=10.1074/jbc.272.42.26479;  
 RA Kast C., Gros P.;  
 RT "Topology mapping of the amino-terminal half of multidrug resistance-  
 RT associated protein by epitope insertion and immunofluorescence.";  
 RL J. Biol. Chem. 272:26479-26487(1997).  
 RN [8]  
 RP TOPOLOGY.  
 RX MEDLINE=98153110; PubMed=9485377; DOI=10.1021/bi97332v;  
 RA Kast C., Gros P.;  
 RT "Epitope insertion favors a six transmembrane domain model for the  
 RT carboxy-terminal portion of the multidrug resistance-associated  
 RT protein.";  
 RL Biochemistry 37:2305-2313(1998).  
 RN [9]  
 RP MUTAGENESIS OF ASP-792; ASP-793; LYS-1333 AND 1454-ASP-GLU-1455.  
 RX MEDLINE=21362977; PubMed=11469806; DOI=10.1006/abbi.2001.2441;  
 RA Cui L., Hou Y.-X., Riordan J.R., Chang X.-B.;  
 RT "Mutations of the Walker B motif in the first nucleotide binding  
 RT domain of multidrug resistance protein MRP1 prevent conformational  
 RT maturation.";  
 RL Arch. Biochem. Biophys. 392:153-161(2001).  
 RN [10]  
 RP MUTAGENESIS OF TRP-1246.  
 RX MEDLINE=21238219; PubMed=11278867; DOI=10.1074/jbc.M011246200;  
 RA Ito K., Olsen S.L., Qiu W., Deeley R.G., Cole S.P.C.;  
 RT "Mutation of a single conserved tryptophan in multidrug resistance  
 RT protein 1 (MRP1/ABCC1) results in loss of drug resistance and  
 RT selective loss of organic anion transport.";  
 RL J. Biol. Chem. 276:15616-15624(2001).  
 RN [11]  
 RP VARIANTS GLN-633 AND VAL-671.  
 RX MEDLINE=20283940; PubMed=10835642; DOI=10.1038/76102;  
 RA Le Saux O., Urban Z., Tschuch C., Csaszar K., Bacchelli B.,  
 RA Quagliari D., Fasquali-Ronchetti I., Pope F.M., Richards A., Terry S.,  
 RA Bercovich L., de Paape A., Boyd C.D.;  
 RT "Mutations in a gene encoding an ABC transporter cause pseudoxanthoma  
 RT elasticum.";  
 RL Nat. Genet. 25:223-227(2000).  
 RN [12]  
 RP VARIANT VAL-671.  
 RX MEDLINE=20283940; PubMed=10835642; DOI=10.1038/76102;  
 RA Ringfeil F., Lebowitz M.G., Cristiano A.M., Uitto J.;  
 RT "Pseudoxanthoma elasticum: mutations in the MRP6 gene encoding a  
 RT transmembrane ATP-binding cassette (ABC) transporter.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 97:6001-6006(2000).  
 RN [13]  
 RP VARIANT SER-433, AND CHARACTERIZATION OF VARIANT VAL-671.  
 RX MEDLINE=21578494; PubMed=11721885;  
 RA Conrad S., Kauffmann H.-M., Ito K., Deeley R.G., Cole S.P.C.,  
 RA Schrenk D.;  
 RT "Identification of human multidrug resistance protein 1 (MRP1)  
 RT mutations and characterization of a G671V substitution.";  
 RL J. Hum. Genet. 46:656-663(2001).  
 RN [14]  
 RP VARIANTS MET-117 AND LEU-1512.  
 RX MEDLINE=20579883; PubMed=11139250;  
 RA DOI=10.1002/1098-1004(2001)17:1<74::AID-HUMU14-3.0.CO;2-F;  
 RA Perdu J., Germain D.P.;  
 RT "Identification of novel polymorphisms in the pM5 and MRP1 (ABCC1)  
 RT genes at locus 16p13.1 and exclusion of both genes as responsible for  
 RT pseudoxanthoma elasticum.";  
 RL Hum. Mutat. 17:74-75(2001).  
 RN [15]  
 RP VARIANTS SER-433; ILE-73; GLN-723 AND GLN-1058.  
 RX MEDLINE=21163948; PubMed=11266082;  
 RA Ito S., Ieiri I., Tanabe M., Suzuki A., Higuchi S., Otsubo K.;  
 RT "Polymorphism of the ABC transporter genes, MDR1, MRP1 and MRP2/CMOAT,  
 RT in healthy Japanese subjects.";  
 RL Pharmacogenetics 11:175-184(2001).  
 RN [16]  
 RP FUNCTION: May participate directly in the active transport of  
 RP drugs into subcellular organelles or influence drug distribution  
 RP indirectly.  
 RP SUBCELLULAR LOCATION: Integral membrane protein.  
 RP ALTERNATIVE PRODUCTS:  
 CC Event-Alternative splicing; Named isoforms=8;  
 CC Comment-Additional isoforms seem to exist. Experimental  
 CC confirmation may be lacking for some isoforms;  
 CC Name=Alleles;  
 CC IsoId=P33527-1; Sequence=Displayed;  
 CC Name=Delexon-17;  
 CC IsoId=P33527-2; Sequence=VSP\_000037;  
 CC Name=Delexon-18;  
 CC IsoId=P33527-3; Sequence=VSP\_000038;  
 CC Name=Delexon-30;  
 CC IsoId=P33527-4; Sequence=VSP\_000039;  
 CC Name=Delexon-17-18;  
 CC IsoId=P33527-5; Sequence=VSP\_000037, VSP\_000038;  
 CC Name=Delexon-17-30;  
 CC IsoId=P33527-6; Sequence=VSP\_000037, VSP\_000039;  
 CC Name=Delexon-18-30;  
 CC IsoId=P33527-7; Sequence=VSP\_000038, VSP\_000039;  
 CC Name=Delexon-17-18-30;  
 CC IsoId=P33527-8; Sequence=VSP\_000037, VSP\_000038, VSP\_000039;  
 CC TISSUE SPECIFICITY: Lung, testis and peripheral blood mononuclear  
 CC cells.  
 CC SIMILARITY: Belongs to the ABC transporter family. MRP subfamily.  
 CC DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;  
 CC WWW=http://www.infobiogen.fr/services/chronocancer/Genes/MRPID106.html".  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL; L05628; AAB46616.1; --  
 DR EMBL; AF022853; AAB83983.1; JOINED.  
 DR EMBL; AF022824; AAB83983.1; JOINED.  
 DR EMBL; AF022825; AAB83983.1; JOINED.  
 DR EMBL; AF022826; AAB83983.1; JOINED.  
 DR EMBL; AF022827; AAB83983.1; JOINED.  
 DR EMBL; AF022828; AAB83983.1; JOINED.  
 DR EMBL; AF022829; AAB83983.1; JOINED.  
 DR EMBL; AF022830; AAB83983.1; JOINED.  
 DR EMBL; AF022831; AAB83983.1; JOINED.  
 DR EMBL; AF022832; AAB83983.1; JOINED.  
 DR EMBL; AF022833; AAB83983.1; JOINED.  
 DR EMBL; AF022834; AAB83983.1; JOINED.  
 DR EMBL; AF022835; AAB83983.1; JOINED.  
 DR EMBL; AF022836; AAB83983.1; JOINED.  
 DR EMBL; AF022837; AAB83983.1; JOINED.  
 DR EMBL; AF022838; AAB83983.1; JOINED.  
 DR EMBL; AF022839; AAB83983.1; JOINED.  
 DR EMBL; AF022840; AAB83983.1; JOINED.  
 DR EMBL; AF022841; AAB83983.1; JOINED.  
 DR EMBL; AF022842; AAB83983.1; JOINED.  
 DR EMBL; AF022843; AAB83983.1; JOINED.  
 DR EMBL; AF022844; AAB83983.1; JOINED.  
 DR EMBL; AF022845; AAB83983.1; JOINED.  
 DR EMBL; AF022846; AAB83983.1; JOINED.  
 DR EMBL; AF022847; AAB83983.1; JOINED.  
 DR EMBL; AF022848; AAB83983.1; JOINED.  
 DR EMBL; AF022849; AAB83983.1; JOINED.  
 DR EMBL; AF022850; AAB83983.1; JOINED.



EMBL; AF022840; AAB83979.1; JOINED.  
 DR EMBL; AF022841; AAB83979.1; JOINED.  
 DR EMBL; AF022842; AAB83979.1; JOINED.  
 DR EMBL; AF022843; AAB83979.1; JOINED.  
 DR EMBL; AF022844; AAB83979.1; JOINED.  
 DR EMBL; AF022845; AAB83979.1; JOINED.  
 DR EMBL; AF022846; AAB83979.1; JOINED.  
 DR EMBL; AF022847; AAB83979.1; JOINED.  
 DR EMBL; AF022848; AAB83979.1; JOINED.  
 DR EMBL; AF022849; AAB83979.1; JOINED.  
 DR EMBL; AF022850; AAB83979.1; JOINED.  
 DR EMBL; AF022851; AAB83979.1; JOINED.  
 DR EMBL; AF022852; AAB83979.1; JOINED.  
 DR EMBL; AF022853; AAB83979.1; JOINED.  
 DR HSSP; P08716; IMT0.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0043626; F:ATPase activity, coupled to transmembrane m. . .; IEA.  
 DR GO; GO:000166; F:nucleotide binding; IEA.  
 DR GO; GO:0005215; F:transporter activity; IEA.  
 DR GO; GO:0006810; P:transport; IEA.  
 DR InterPro; IPR003593; AAA ATPase.  
 DR InterPro; IPR011527; ABC membrane 1.  
 DR InterPro; IPR001140; ABC TM transp.  
 DR InterPro; IPR003439; ABC transporter.  
 DR InterPro; IPR005292; MRP\_assoc.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR Pfam; PF00664; ABC membrane; 2.  
 DR Pfam; PF00005; ABC tran; 2.  
 DR ProDom; PD000006; ABC transporter; 2.  
 DR SMART; SM00382; AAA; 2.  
 DR TIGRFAM; TIGR00957; MRP assoc\_pro; 1.  
 DR PROSITE; PS00929; ABC\_TMIF; 2.  
 DR PROSITE; PS00211; ABC\_TRANSPORTER\_1; 2.  
 DR PROSITE; PS00893; ABC\_TRANSPORTER\_2; 2.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; UNKNOWN\_1.  
 KW ATP-binding.  
 FT NON\_TER  
 SQ SEQUENCE 1 169852 MW; DC85592817C439FE CRC64;  
 Query Match 77.6%; Score 7769; DB 2; Length 1515;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1515; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 17 DNNVTWNTSNPDFTKCFQNTLVWVPCFYLWACFPFFFLYLSRHRDGYIQMTPLNKTKA 76  
 DB 1 DNNVTWNTSNPDFTKCFQNTLVWVPCFYLWACFPFFFLYLSRHRDGYIQMTPLNKTKA 60  
 QY 77 LGFLWLVWVWADLFYSFWRSGRIFLAPVFLVSPPTLIGITTLATFLIQLERRKGVQSSG 136  
 DB 61 LGFLWLVWVWADLFYSFWRSGRIFLAPVFLVSPPTLIGITTLATFLIQLERRKGVQSSG 120  
 QY 137 IMLTFLWLVWVWADLFYSFWRSGRIFLAPVFLVSPPTLIGITTLATFLIQLERRKGVQSSG 196  
 DB 121 IMLTFLWLVWVWADLFYSFWRSGRIFLAPVFLVSPPTLIGITTLATFLIQLERRKGVQSSG 180  
 QY 197 LFSETHDNPCESSASFLSRITFWITGLIVRGYRQPLEGSLWLNKEDTSEQVVPV 256  
 DB 181 LFSETHDNPCESSASFLSRITFWITGLIVRGYRQPLEGSLWLNKEDTSEQVVPV 240  
 QY 257 LVXNWKKECAKTRQPKVWVYSSKDPAPQKESKSKVDANEVEALIVKSPQKEMNPSLFKV 316  
 DB 241 LVXNWKKECAKTRQPKVWVYSSKDPAPQKESKSKVDANEVEALIVKSPQKEMNPSLFKV 300  
 QY 317 LYKTPGPFLMSFFPKAIDHLMFSGPQILKLIKFNVDTKAPDQGYFTVLLFVTAQL 376  
 DB 301 LYKTPGPFLMSFFPKAIDHLMFSGPQILKLIKFNVDTKAPDQGYFTVLLFVTAQL 360  
 QY 377 QTLVLHQYFHCIFVSGMEIKTAVTGAVYRKALVITNSARKSTVGEIVNLMSVDAQRFMD 436

DB 361 QTLVLHQYFHCIFVSGMEIKTAVTGAVYRKALVITNSARKSTVGEIVNLMSVDAQRFMD 420  
 QY 437 LATYINMIWSAPLQVILALYLLNLGSPSVLAGVAVMVLMPVNVNMAWMTKTYQVAHMK 496  
 DB 421 LATYINMIWSAPLQVILALYLLNLGSPSVLAGVAVMVLMPVNVNMAWMTKTYQVAHMK 480  
 QY 497 SKDNRIKLMMNEILNGIKVLYAYWELAFKDKVLAIRQBELKVLKKSAYLSAVGFTTWCT 556  
 DB 481 SKDNRIKLMMNEILNGIKVLYAYWELAFKDKVLAIRQBELKVLKKSAYLSAVGFTTWCT 540  
 QY 557 PFLVALCTFAVVYITDENNILDQAOTAFVSLALFNILREPLNLPWVSIIVQASVSLKRL 616  
 DB 541 PFLVALCTFAVVYITDENNILDQAOTAFVSLALFNILREPLNLPWVSIIVQASVSLKRL 600  
 QY 617 RIFLSHELEPDSIERRPVKDGGGTNSITVRNATFTWARSDDPTINGITTFISPEGALVAV 676  
 DB 601 RIFLSHELEPDSIERRPVKDGGGTNSITVRNATFTWARSDDPTINGITTFISPEGALVAV 660  
 QY 677 VGQVCGCKSSLLSALLAEMDKVEGHVAIKGSVAVVPOQAWIQNDSLRENILFGCOLREPY 736  
 DB 661 VGQVCGCKSSLLSALLAEMDKVEGHVAIKGSVAVVPOQAWIQNDSLRENILFGCOLREPY 720  
 QY 737 YRSVIQACALLPDLLEILPSPGDRTEIGEGKGNLSGGQKQKQVSLARAVYNSADIYLFDDPLS 796  
 DB 721 YRSVIQACALLPDLLEILPSPGDRTEIGEGKGNLSGGQKQKQVSLARAVYNSADIYLFDDPLS 780  
 QY 797 AVDAHVGKHFENVIGPKGMLKNKTRILVTHSMYSLPQVDVIIVMGSGKISEMGSYQBL 856  
 DB 781 AVDAHVGKHFENVIGPKGMLKNKTRILVTHSMYSLPQVDVIIVMGSGKISEMGSYQBL 840  
 QY 857 ARDGAFAEFLRTYASTEODABENGVTGSGPCKEAKOMENGLVTDTSAGKQOROLSS 916  
 DB 841 ARDGAFAEFLRTYASTEODABENGVTGSGPCKEAKOMENGLVTDTSAGKQOROLSS 900  
 QY 917 SSSYSGDISRHHNSTABEQKAEAKKEETWKLMEADKATQGVKLSVYMDYKAIGLFISF 976  
 DB 901 SSSYSGDISRHHNSTABEQKAEAKKEETWKLMEADKATQGVKLSVYMDYKAIGLFISF 960  
 QY 977 LSIFLFCNHNVSALASNYWLSLWTDPTVNGTQHTKVRSLVYCALGISQIAVFGYSMA 1036  
 DB 961 LSIFLFCNHNVSALASNYWLSLWTDPTVNGTQHTKVRSLVYCALGISQIAVFGYSMA 1020  
 QY 1037 VSIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFSKELTDVDSMPEIKMFMG 1096  
 DB 1021 VSIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFSKELTDVDSMPEIKMFMG 1080  
 QY 1097 SLFNVIACIVILLATPIAAIIIPPLGLIYFFVQRFVYVASSRQLKRLESVSRSPVYSHFN 1156  
 DB 1081 SLFNVIACIVILLATPIAAIIIPPLGLIYFFVQRFVYVASSRQLKRLESVSRSPVYSHFN 1140  
 QY 1157 ETLGVSIVIRAFERQERFIHQSDLKVDENOKAYYPSIVANRWLAVRLECVCNLCVLAAL 1216  
 DB 1141 ETLGVSIVIRAFERQERFIHQSDLKVDENOKAYYPSIVANRWLAVRLECVCNLCVLAAL 1200  
 QY 1217 FAVISRHSLSAGLVGLSVSLQVTTYINLWLRMSSEMETNIIVAVERLKEYSETKEAPW 1276  
 DB 1201 FAVISRHSLSAGLVGLSVSLQVTTYINLWLRMSSEMETNIIVAVERLKEYSETKEAPW 1260  
 QY 1277 QIQTAPSSWPQVGRVEFRNYCLRYREDLDVLRHINVTINGEKVIGVIGRTGAGKSSL 1336  
 DB 1261 QIQTAPSSWPQVGRVEFRNYCLRYREDLDVLRHINVTINGEKVIGVIGRTGAGKSSL 1320  
 QY 1337 TLGLFRINESAEGEIIIDGINIAKIGLHDLRFKTTIIPQDPVLFGSLRNLPFPFSQYSD 1396  
 DB 1321 TLGLFRINESAEGEIIIDGINIAKIGLHDLRFKTTIIPQDPVLFGSLRNLPFPFSQYSD 1380  
 QY 1397 EEWVTSLELAHLKDFVSALPDKLDECAEGENLSVQROLVCLARALLKTKILVLEA 1456  
 DB 1381 EEWVTSLELAHLKDFVSALPDKLDECAEGENLSVQROLVCLARALLKTKILVLEA 1440  
 QY 1457 TAAVDLETDDLIOSTIRTOQEDCTVLTIAHRLNTIMDYTRVIVLDKGEIOEYGAPSDLLQ 1516  
 DB 1441 TAAVDLETDDLIOSTIRTOQEDCTVLTIAHRLNTIMDYTRVIVLDKGEIOEYGAPSDLLQ 1500

```

QY 1517 QRLFLYSMAKDAGLV 1531
DB 1501 QRLFLYSMAKDAGLV 1515

RESULT 3
Q864R9
AC Q864R9 PRELIMINARY; PRT; 1531 AA.
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Multidrug resistance protein 1B.
GN Name=MRP1;
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopithecinæ; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22544876; PubMed=12657726;
RA Codinot N., Iversen P.W., Tabas L., Xia X., Williams D.C.,
RA Dantzig A.H., Perry W.L.;
RT "Cloning and functional characterization of the multidrug resistance-
RT associated protein (MRP1/ABCC1) from the cynomolgus monkey.";
RL Mol. Cancer Ther. 2:307-316(2003).
RN [2]
RP SEQUENCE FROM N.A.
RA Perry W.L. III, Godinot N.;
RA Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AY146673; AAN65349.1; -
DR HSP; P08716; 1MT0.
DR GO; GO:0015021; C:integral to membrane; IEA.
DR GO; GO:0005522; F:ATP binding; IEA.
DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR011527; ABC_membrane_1.
DR InterPro; IPR001140; ABC_TM_transp.
DR InterPro; IPR003439; ABC_transporter.
DR InterPro; IPR005292; MRP_assoc.
DR InterPro; IPR000719; Prot_kinase.
DR Pfam; PF00664; ABC_membrane; 2.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC_transporter; 2.
DR SMART; SM00382; AAA; 2.
DR TIGRFAMs; TIGR00957; MRP_assoc_pro; 1.
DR PROSITE; PS50929; ABC_TMIF; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS50893; ABC_TRANSPORTER_2; 2.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
KW ATP-binding.
SQ SEQUENCE 1531 AA; 171658 MW; 1AE788BFD9EF459 CRC64;

Query Match 76.8%; Score 7695; DB 2; Length 1531;
Best Local Similarity 97.5%; Pred. No. 0;
Matches 1493; Conservative 20; Mismatches 18; Indels 0; Gaps 0;

QY 1 MALRGFCSADGSDPLWDMNVNTSNPDKFQNTLVWVPCFYLWACFPFYLYLSRH 60
DB 1 MALRGFCSADGSDPLWDMNVNTSNPDKFQNTLVWVPCFYLWACFPFYLYLSRH 60
QY 61 DRGYIQMTPLNKTALGFLMIVCWADLFYFWSRSGIFLAPFLVSPITLIGITLLIA 120
DB 61 DRGYIQMTPLNKTALGFLMIVCWADLFYFWSRSGIFLAPFLVSPITLIGITLLIA 120
QY 121 TFLIQLERRKGVQSSGIMLTFWLVALVCAILIRSKIMTALKEDQVDLFRDITFYVPS 180
DB 121 TFLIQLERRKGVQSSGIMLTFWLVALVCAILIRSKIMTALKEDQVDLFRDITFYVPS 180

```

QY 1261 VERLKEYSETEAPWQIETAPPSSWPOGVERFNRYCLRYREDLDFVLRHINTVING 1320  
 DB 1261 VERLKEYSETEAPWQIETAPPSSWPOGVERFNRYCLRYREDLDFVLRHINTVING 1320  
 QY 1321 EKVGIVGRTGAGKSSITLGLPRINSAEGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
 DB 1321 EKVGIVGRTGAGKSSITLGLPRINSAEGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
 QY 1381 SGLSRNLDPPSOYSDSEEVWTSLELAHLKDFYSALPKLDHCEAGGENLSVGQRLVCL 1440  
 DB 1381 SGLSRNLDPPSOYSDSEEVWTSLELAHLKDFYSALPKLDHCEAGGENLSVGQRLVCL 1440  
 QY 1441 ARALLRKTILVLDATAVLDLTDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 DB 1441 ARALLRKTILVLDATAVLDLTDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 QY 1501 DXGEIQEYCAPSDLLQORGLFYSMAKDAGLV 1531  
 DB 1501 DXGEIQEYCAPSDLLQORGLFYNMARDAGLV 1531

## RESULT 4

Q864S0 PRELIMINARY; PRT; 1531 AA.  
 AC Q864S0;  
 DT 01-JUN-2003 (TReMBLrel. 24, Created)  
 DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)  
 DT 01-WAR-2004 (TReMBLrel. 26, Last annotation update)  
 DE Multidrug resistance protein 1A.  
 GN Name=MRP1;  
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;  
 OC Cercopitheciinae; Macaca.  
 OX NCBI\_TaxId=9541;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RX MEDLINE=22544876; PubMed=12657726;  
 RA Godinot N., Iversen P.W., Tabas L., Xia X., Williams D.C.,  
 RA Dantzig A.H., Perry W.L.;  
 RT "Cloning and functional characterization of the multidrug resistance-  
 RT associated protein (MRP1/ABCC1) from the cynomolgus monkey.";  
 RL Mol. Cancer Ther. 2:307-316(2003).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RA Perry W.L., Ill, Godinot N.;  
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBSJ databases.  
 CC -I- SIMILARITY: Belongs to the ABC transporter family.  
 DR EMBL; AY146672; AAN65348.1; -;  
 DR HSSP; P08716; 1MT0.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.  
 DR GO; GO:0000166; F:nucleotide binding; IEA.  
 DR GO; GO:0005215; F:transporter activity; IEA.  
 DR GO; GO:0006810; P:transporter; IEA.  
 DR InterPro; IPR003593; AAA\_ATPase.  
 DR InterPro; IPR011527; ABC\_membrane\_1.  
 DR InterPro; IPR001140; ABC\_TM\_transpt.  
 DR InterPro; IPR003439; ABC\_transporter.  
 DR InterPro; IPR005292; MRP\_assoc.  
 DR InterPro; IPR000719; Prot\_Kinase.  
 DR Pfam; PF00664; ABC\_membrane; 2.  
 DR Pfam; PF00005; ABC\_tran; 2.  
 DR Prodom; PD00006; ABC\_transporter; 2.  
 DR SMART; SM00382; AAA; 2.  
 DR TIGRFAMs; TIGR00957; MRP\_assoc\_pro; 1.  
 DR PROSITE; PS50929; ABC\_TMIF; 2.  
 DR PROSITE; PS00211; ABC\_TRANSPORTER\_1; 2.  
 DR PROSITE; PS50893; ABC\_TRANSPORTER\_2; 2.

DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; UNKNOWN\_1.  
 KW ATP-binding.  
 SQ SEQUENCE 1531 AA; 171700 MW; 168712B8CC2D2B89 CRC64;  
 Query Match 76.8%; Score 7689; DB 2; Length 1531;  
 Best Local Similarity 97.6%; Pred. No. 0;  
 Matches 1494; Conservative 18; Mismatches 19; Indels 0; Gaps 0;  
 QY 1 MALRGFCSADSDPLMDWNVNTWNTSNPDFTKCFQNTVLVWVPCFYLMACFFPYFLYLSRH 60  
 DB 1 MALRGFCSADSDPLMDWNVNTWNTSNPDFTKCFQNTVLVWVPCFYLMACFFPYFLYLSRH 60  
 QY 61 DRGVIQMTPLNKTALGFLIWCWADLFYSFERSRGIFLAPVFLVSPITLLGITLLA 120  
 DB 61 DRGVIQMTPLNKTALGFLIWCWADLFYSFERSRGIFLAPVFLVSPITLLGITLLA 120  
 QY 121 TFLIQLERRKGVQSSGIMLTFWLVALVCAALILRSKIMTALKEDAQVDLFRDITFYVYFS 180  
 DB 121 TFLIQLERRKGVQSSGIMLTFWLVALVCAALILRSKIMTALKEDAQVDLFRDITFYVYFS 180  
 QY 181 LLLIQLVLSCESDRSPISFSETHDNPCESSASFLSRITFWWITGLIVRGYRQPLEGSD 240  
 DB 181 LLLIQLVLSCESDRSPISFSETHDNPCESSASFLSRITFWWITGLIVRGYRQPLEGSD 240  
 QY 241 LWSLNKEDTSEQVVPVLVKNWKKCAKTRKQPVVVVSSKDPAPQPKSSKVDANEEVEAL 300  
 DB 241 LWSLNKEDTSEQVVPVLVKNWKKCAKTRKQPVVVVSSKDPAPQPKSSKVDANEEVEAL 300  
 QY 301 IVKSPQKWNPSLPKVLKTPGYPFLMSFPFKAHDLMMFSGPOILKLLIKFVNDTRAPD 360  
 DB 301 IVKSPQKWNPSLPKVLKTPGYPFLMSFPFKAHDLMMFSGPOILKLLIKFVNDTRAPD 360  
 QY 361 WQGYFYTVLLFVTAACLOTVLHGVFHCYFGSMRIKTAIVGAVRKALVITNSARKSTV 420  
 DB 361 WQGYFYTVLLFVTAACLOTVLHGVFHCYFGSMRIKTAIVGAVRKALVITNSARKSTV 420  
 QY 421 GEIVNLSVDAQREFMDLATYINMIWSAPLOVILALYLLNMLNLSVLAGVAVMLVMPVN 480  
 DB 421 GEIVNLSVDAQREFMDLATYINMIWSAPLOVILALYLLNMLNLSVLAGVAVMLVMPVN 480  
 QY 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAWELAFKDKVLAIROBELKVLK 540  
 DB 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAWELAFKDKVLAIROBELKVLK 540  
 QY 541 KSAVLSAVGTFTWVCTPFLVALCTFVYVITDENILDAOTAFVLSLAFNLIRPLNLIP 600  
 DB 541 KSAVLSAVGTFTWVCTPFLVALCTFVYVITDENILDAOTAFVLSLAFNLIRPLNLIP 600  
 QY 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWARSDDPT 660  
 DB 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWARSDDPT 660  
 QY 661 LNGITFSIPREGALVAVVGVQCGKSSLLSALLAEMDKVEGHVAIKGSVAVYPOQAWTND 720  
 DB 661 LNGITFSIPREGALVAVVGVQCGKSSLLSALLAEMDKVEGHVAIKGSVAVYPOQAWTND 720  
 QY 721 SLRENILFGQLEPEPYRVSIVIOACALLPDLIELPSGDRTEIGEGVNLSGQKORVSLAR 780  
 DB 721 SLRENILFGQLEPEPYRVSIVIOACALLPDLIELPSGDRTEIGEGVNLSGQKORVSLAR 780  
 QY 781 AVYSNADIYLFDDPPLSAVDAHVGKHFENVIGPKGMLKNKTRILVTHSMYSYLPQVDVIIV 840  
 DB 781 AVYSNADIYLFDDPPLSAVDAHVGKHFENVIGPKGMLKNKTRILVTHSMYSYLPQVDVIIV 840  
 QY 841 MSGGKISEMGSYQELLARDGAFAEFLRTYASTEQDAEENGVTGVSQKGAEMNGM 900  
 DB 841 MSGGKISEMGSYQELLARDGAFAEFLRTYASTEQDAEENGVTGVSQKGAEMNGM 900  
 QY 901 LVTDGAKQLQRLSSSSSYSGDISRHHNSTAELOKAEAKKEETWKLMEADKATQGVKL 960  
 DB 901 LVTDGAKQLQRLSSSSSYSGDISRHHNSTAELOKAEAKKEETWKLMEADKATQGVKL 960  
 QY 961 SVYWDYMKAILGLFISLISIFLMCNHVSALASNYWLSLTDTPDIPVNGTQEHKVRLLSYG 1020

Db 961 SVYDYMKAIGLPSLSIFLFTCNHVAALASNYLWLSWTDDIIVNGTOEHTKVRLSVYG 1020  
 QY 1021 ALGISOGIAVFGYSMAVSGIGILASRCHLVLDLHLSILRSPMSFFERTPPSNLVNRFKSKEL 1080  
 Db 1021 ALGISOGIAVFGYSMAVSGIGILASRCHLVLDLHLSILRSPMSFFERTPPSNLVNRFKSKEL 1080  
 QY 1081 DTVDMSIPEVIMKFMGSLFNVI GACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQL 1140  
 Db 1081 DTVDMSIPEVIMKFMGSLFNVI GACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQL 1140  
 QY 1141 KRLESVRSRVPYSHFNETLLGVSVIRAFERQERFIHQSDLKVDENOKAYPSIVANRWLA 1200  
 Db 1141 KRLESVRSRVPYSHFNETLLGVSVIRAFERQERFIHQSDLKVDENOKAYPSIVANRWLA 1200  
 QY 1201 VRLECVGNCIVLFAALFAVISRSLSGAGLVGLSVSYSLQVTTVNLNLRMSSEMETNIVA 1260  
 Db 1201 VRLECVGNCIVLFAALFAVISRSLSGAGLVGLSVSYSLQVTTVNLNLRMSSEMETNIVA 1260  
 QY 1261 VERLKEYSETEKAPWQIQETAPSSWPQVGRVEFRNYCLRYREDLDFVLRHINVINGG 1320  
 Db 1261 VERLKEYSETEKAPWQIQETAPSSWPQVGRVEFRNYCLRYREDLDFVLRHINVINGG 1320  
 QY 1321 EKVGIVGRTGAGKSSLTGLFRINESAEGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
 Db 1321 EKVGIVGRTGAGKSSLTGLFRINESAEGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
 QY 1381 SGLRNLNLPFSOYDEEVTSLLELAHLKDFVSALPKLDHCEAEGENLSVQROLVCL 1440  
 Db 1381 SGLRNLNLPFSOYDEEVTSLLELAHLKDFVSALPKLDHCEAEGENLSVQROLVCL 1440  
 QY 1441 ARALLKTKILVLDEATAAVDLTDDLIQSTIRTFQEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 Db 1441 ARALLKTKILVLDEATAAVDLTDDLIQSTIRTFQEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 QY 1501 DKGEIQEYGAPSDLLQOQGLFYFMAKDAGLV 1531  
 Db 1501 DKGEIQEYGAPSDLLQOQGLFYFMAKDAGLV 1531

RESULT 5

Q9UQ97 PRELIMINARY; PRT; 1459 AA.  
 ID Q9UQ97  
 AC Q9UQ97  
 DT 01-MAY-2000 (T-EMBLrel. 13, Created)  
 DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)  
 DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)  
 DE Multidrug resistance protein (fragment).  
 GN Names=MRP;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98008927; PubMed=9344662; DOI=10.1006/geno.1997.4950;  
 RA Grant C.E., Kurz E.U., Cole S.P., Deeley R.G.;  
 RT "Analysis of the intron-exon organization of the human multidrug-  
 RT resistance protein gene (MRP) and alternative splicing of its mRNA";  
 RL Genomics 45:368-378 (1997).  
 CC -!- SIMILARITY: Belongs to the ABC transporter family.  
 DR EMBL; AF022824; AAB83980.1; JOINED.  
 DR EMBL; AF022825; AAB83980.1; JOINED.  
 DR EMBL; AF022826; AAB83980.1; JOINED.  
 DR EMBL; AF022827; AAB83980.1; JOINED.  
 DR EMBL; AF022828; AAB83980.1; JOINED.  
 DR EMBL; AF022829; AAB83980.1; JOINED.  
 DR EMBL; AF022830; AAB83980.1; JOINED.  
 DR EMBL; AF022831; AAB83980.1; JOINED.  
 DR EMBL; AF022832; AAB83980.1; JOINED.  
 DR EMBL; AF022833; AAB83980.1; JOINED.  
 DR EMBL; AF022834; AAB83980.1; JOINED.  
 DR EMBL; AF022835; AAB83980.1; JOINED.

DR EMBL; AF022836; AAB83980.1; JOINED.  
 DR EMBL; AF022837; AAB83980.1; JOINED.  
 DR EMBL; AF022838; AAB83980.1; JOINED.  
 DR EMBL; AF022839; AAB83980.1; JOINED.  
 DR EMBL; AF022841; AAB83980.1; JOINED.  
 DR EMBL; AF022842; AAB83980.1; JOINED.  
 DR EMBL; AF022843; AAB83980.1; JOINED.  
 DR EMBL; AF022844; AAB83980.1; JOINED.  
 DR EMBL; AF022845; AAB83980.1; JOINED.  
 DR EMBL; AF022846; AAB83980.1; JOINED.  
 DR EMBL; AF022847; AAB83980.1; JOINED.  
 DR EMBL; AF022848; AAB83980.1; JOINED.  
 DR EMBL; AF022849; AAB83980.1; JOINED.  
 DR EMBL; AF022850; AAB83980.1; JOINED.  
 DR EMBL; AF022851; AAB83980.1; JOINED.  
 DR EMBL; AF022852; AAB83980.1; JOINED.  
 DR EMBL; AF022853; AAB83980.1; JOINED.  
 DR HSSP; P08716; IMT0.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0043262; F:ATPase activity, coupled to transmembrane m...; IEA.  
 DR GO; GO:0000166; F:nucleotide binding; IEA.  
 DR GO; GO:0005215; F:transporter activity; IEA.  
 DR GO; GO:0006810; P:transport; IEA.  
 DR InterPro; IPR003593; AAA ATPase.  
 DR InterPro; IPR001140; ABC membrane 1.  
 DR InterPro; IPR001157; ABC TM transp.  
 DR InterPro; IPR003439; ABC transporter.  
 DR InterPro; IPR005292; MRP\_assoc.  
 DR Pfam; PF00664; ABC membrane; 2.  
 DR Pfam; PF00005; ABC tran; 1.  
 DR ProDom; PD000006; ABC transporter; 1.  
 DR SMART; SM00382; AAA; 2.  
 DR TIGRFAMs; TIGR00957; MRP\_assoc\_pro; 1.  
 DR PROSITE; PS00929; ABC TMIF; 2.  
 DR PROSITE; PS00211; ABC TRANSPORTER 1; UNKNOWN 1.  
 DR PROSITE; PS00893; ABC TRANSPORTER 2; 2.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; UNKNOWN 1.  
 KW ATP-binding.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 1459 AA; 163830 MM; 6A016A5A6E7AA95C CRC64;  
 Query Match 74.3%; Score 7444; DB 2; Length 1459;  
 Best Local Similarity 96.3%; Pred. No. 0;  
 Matches 1459; Conservative 0; Mismatches 0; Indels 56; Gaps 1;  
 QY 17 DNNVTWNTSNPDFTKCFQNTLVVWVPCFYLWACFPFYLILSRHGRGYIOMTPLNKTKA 76  
 Db 1 DNNVTWNTSNPDFTKCFQNTLVVWVPCFYLWACFPFYLILSRHGRGYIOMTPLNKTKA 60  
 QY 77 LGFLMWVWADLFYSFWSRSGIFLAPVFLVSPTLGTTLLATLFIQLERRKGVQSSG 136  
 Db 61 LGFLMWVWADLFYSFWSRSGIFLAPVFLVSPTLGTTLLATLFIQLERRKGVQSSG 120  
 QY 137 IMLTFWLVALVCALAILRSKIMTALKEDAQVDFRIDTFYVYFSLILLIQLVSCFSDRSP 196  
 Db 121 IMLTFWLVALVCALAILRSKIMTALKEDAQVDFRIDTFYVYFSLILLIQLVSCFSDRSP 180  
 QY 197 LFSETHDNPCCPSSASFLSRITFWITGLVGRYQPLEGSDMLSNKEDTSEQVVPV 256  
 Db 181 LFSETHDNPCCPSSASFLSRITFWITGLVGRYQPLEGSDMLSNKEDTSEQVVPV 240  
 QY 257 LVKNWKKCEAKTRKQPVKVYSSKDPAPQKESKVDANEVEALIVKSPQKWNPSLFKV 316  
 Db 241 LVKNWKKCEAKTRKQPVKVYSSKDPAPQKESKVDANEVEALIVKSPQKWNPSLFKV 300  
 QY 317 LYKTFGPVFLMSFFFKAIHDLNMFSGPQLIKLIFVNDTKAPDWQGYFYTVLLFVTACL 376  
 Db 301 LYKTFGPVFLMSFFFKAIHDLNMFSGPQLIKLIFVNDTKAPDWQGYFYTVLLFVTACL 360  
 QY 377 QTLVLHQYFHCIVSGMRITKTAIVGAYVRKALVITNSARKSVTGVISVNLMSVDAQRFMD 436

Db 361 QTLVLHQYFHCIFVSGMRKTAIVAGVYRKALVITNSARKSSTVGEIVNLMSVDAQRFMD 420  
Qy 437 LATYINMWSAPLOVITLALYLILWNLGSPVSLAGVAVMLVPPVNAVMAKTKTYQVAHMK 496  
Db 421 LATYINMWSAPLOVITLALYLILWNLGSPVSLAGVAVMLVPPVNAVMAKTKTYQVAHMK 480  
Qy 497 SKDNRIKLMNEILNGIKVLKYLAWELAFKDKVLAIROBELKVLKKSAYLSAVGTFTWVCT 556  
Db 481 SKDNRIKLMNEILNGIKVLKYLAWELAFKDKVLAIROBELKVLKKSAYLSAVGTFTWVCT 540  
Qy 557 PELVALCTFAVYVTTDENNDIDAQTAQFVSLALFNILRPNLPLNPVSIQVQASVSLKRL 616  
Db 541 PELVALCTFAVYVTTDENNDIDAQTAQFVSLALFNILRPNLPLNPVSIQVQASVSLKRL 600  
Qy 617 RIFLSHEELEPDSIERPPVKDGGGNTSITVRNATFTWARSDDPTTLNGITFTSIPGALVAV 676  
Db 601 RIFLSHEELEPDSIERPPVKDGGGNTSITVRNATFTWARSDDPTTLNGITFTSIPGALVAV 660  
Qy 677 VQVCGKSSLLSALLAEMDKVEGHVAKGSVAVYVPOQAWIONDSIRENILFGCOLEBPY 736  
Db 661 VQVCGKSSLLSALLAEMDKVEGHVAKGSVAVYVPOQAWIONDSIRENILFGCOLEBPY 720  
Qy 737 YRSVIOACALLPDLBELPSGDRTEICEKGWNLGGQKQKVSLARAVYSNADIYLFDDPLS 796  
Db 721 YRSVIOACALLPDLBELPSGDRTEICEKGWNLGGQKQKVSLARAVYSNADIYLFDDPLS 747  
Qy 797 AYDAHVGKHI FENVIPGKMLKNKTRILVTHSMVLPQVDVILVMSGGKISEMGSYOELL 856  
Db 748 -----KTRILVTHSMVLPQVDVILVMSGGKISEMGSYOELL 784  
Qy 857 ARDGAFAEFLRYASTEOBQDAENGVTGSPGKEAKQKMGMLVTDVSAGKOLQRLSS 916  
Db 785 ARDGAFAEFLRYASTEOBQDAENGVTGSPGKEAKQKMGMLVTDVSAGKOLQRLSS 844  
Qy 917 SSSYSGDISRHNSHTAELOKAEKKEETWKLMEADKAQGTQVKLSVYWDYMKAIGLFISF 976  
Db 845 SSSYSGDISRHNSHTAELOKAEKKEETWKLMEADKAQGTQVKLSVYWDYMKAIGLFISF 904  
Qy 977 LSIFLPMCNHVSALASNYLSLWTDPIVNGTQHTKVLRSVYGALGISQGLAVFGYSA 1036  
Db 905 LSIFLPMCNHVSALASNYLSLWTDPIVNGTQHTKVLRSVYGALGISQGLAVFGYSA 964  
Qy 1037 VSIGGILASRLCHVDLLHLSILRSPMSFFERTPSGNLVNRFKXELDTVDSMPEVKMFPG 1096  
Db 965 VSIGGILASRLCHVDLLHLSILRSPMSFFERTPSGNLVNRFKXELDTVDSMPEVKMFPG 1024  
Qy 1097 SLFNVIGACIVILLATPIAAIIIPPLGLIYFFQVRYVASSRQLKRLSVSRSPVYSHFN 1156  
Db 1025 SLFNVIGACIVILLATPIAAIIIPPLGLIYFFQVRYVASSRQLKRLSVSRSPVYSHFN 1084  
Qy 1157 ETLGLSVIRAFEEQERFTHQSDLKVDENOKAYPSIVANRWLAVRLECVGNCIVLFAAL 1216  
Db 1085 ETLGLSVIRAFEEQERFTHQSDLKVDENOKAYPSIVANRWLAVRLECVGNCIVLFAAL 1144  
Qy 1217 FAVISRHLSAGLVLGSLYSYSQVTTYLNWLRMSSEMETNIVAVERLKEYSETEKEAPW 1276  
Db 1145 FAVISRHLSAGLVLGSLYSYSQVTTYLNWLRMSSEMETNIVAVERLKEYSETEKEAPW 1204  
Qy 1277 QIQTAPPSSWPQGVKEPRNYCLARYREDLPVLRHINVTNGEKGKVGIVGRTGAKGSSL 1336  
Db 1205 QIQTAPPSSWPQGVKEPRNYCLARYREDLPVLRHINVTNGEKGKVGIVGRTGAKGSSL 1264  
Qy 1337 TLGLFRINESAGEIIIDGINAKTGLHDLRPKITIIPODPVLFSGSLRMNLDPPSQYSD 1396  
Db 1265 TLGLFRINESAGEIIIDGINAKTGLHDLRPKITIIPODPVLFSGSLRMNLDPPSQYSD 1324  
Qy 1397 EEVWTSLELAHLKDFVSALPDKLDHECAEGGNSLVGQQLVCLARALRKTKILVLDEA 1456  
Db 1325 EEVWTSLELAHLKDFVSALPDKLDHECAEGGNSLVGQQLVCLARALRKTKILVLDEA 1384  
Qy 1457 TAAVDLETDDLQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVLQKEIQEYGAQSDLLQ 1516  
Db 1385 TAAVDLETDDLQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVLQKEIQEYGAQSDLLQ 1444

Qy 1517 ORGLFYSMAXDAGLV 1531  
Db 1445 ORGLFYSMAXDAGLV 1459

## RESULT 6

OSUQA0 PRELIMINARY; PRT; 1456 AA.

AC Q9UQA0;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
DE Multidrug resistance protein (Fragment).  
CN Name=MRP;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98008927; PubMed=9344652; DOI=10.1006/geno.1997.4950;  
RA Grant C.E., Kurz E.U., Cole S.P., Deeley R.G.;  
RT "Analysis of the intron-exon organization of the human multidrug-  
RT resistance protein gene (MRP) and alternative splicing of its mRNA.";  
RL Genomics 45:368-378(1997).  
CC 1- SIMILARITY: Belongs to the ABC transporter family.  
DR EMBL; AF022824; AAB83981.1; JOINED.  
DR EMBL; AF022825; AAB83981.1; JOINED.  
DR EMBL; AF022826; AAB83981.1; JOINED.  
DR EMBL; AF022827; AAB83981.1; JOINED.  
DR EMBL; AF022828; AAB83981.1; JOINED.  
DR EMBL; AF022829; AAB83981.1; JOINED.  
DR EMBL; AF022830; AAB83981.1; JOINED.  
DR EMBL; AF022831; AAB83981.1; JOINED.  
DR EMBL; AF022832; AAB83981.1; JOINED.  
DR EMBL; AF022833; AAB83981.1; JOINED.  
DR EMBL; AF022834; AAB83981.1; JOINED.  
DR EMBL; AF022835; AAB83981.1; JOINED.  
DR EMBL; AF022836; AAB83981.1; JOINED.  
DR EMBL; AF022837; AAB83981.1; JOINED.  
DR EMBL; AF022838; AAB83981.1; JOINED.  
DR EMBL; AF022840; AAB83981.1; JOINED.  
DR EMBL; AF022841; AAB83981.1; JOINED.  
DR EMBL; AF022842; AAB83981.1; JOINED.  
DR EMBL; AF022843; AAB83981.1; JOINED.  
DR EMBL; AF022844; AAB83981.1; JOINED.  
DR EMBL; AF022845; AAB83981.1; JOINED.  
DR EMBL; AF022846; AAB83981.1; JOINED.  
DR EMBL; AF022847; AAB83981.1; JOINED.  
DR EMBL; AF022848; AAB83981.1; JOINED.  
DR EMBL; AF022849; AAB83981.1; JOINED.  
DR EMBL; AF022850; AAB83981.1; JOINED.  
DR EMBL; AF022851; AAB83981.1; JOINED.  
DR EMBL; AF022852; AAB83981.1; JOINED.  
DR EMBL; AF022853; AAB83981.1; JOINED.  
DR HSSP; P08716; 1M70.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.  
DR GO; GO:0000166; F:nucleotide binding; IEA.  
DR GO; GO:0005215; F:transporter activity; IEA.  
DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR003593; AAA ATPase.  
DR InterPro; IPR011527; ABC membrane\_1.  
DR InterPro; IPR003439; ABC TM transpt.  
DR InterPro; IPR003439; ABC transporter.  
DR InterPro; IPR005292; MRP assoc.  
DR InterPro; IPR000719; Prot Kinase.  
DR Pfam; PF00664; ABC membrane; 2.  
DR Pfam; PF00005; ABC\_tran; 2.  
DR ProDom; PD000006; ABC\_transporter; 2.  
DR SMART; SM00382; AAA; 2.

DR TIGR00957: MRP assoc\_pro; 1.  
 DR PROSITE; PS50929; ABC\_TMIP; 2.  
 DR PROSITE; PS00211; ABC\_TRANSPORTER\_1; 2.  
 DR PROSITE; PS00893; ABC\_TRANSPORTER\_2; 2.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; UNKNOWN\_1.  
 KW ATP-binding.  
 FT NON\_TER  
 SQ SEQUENCE 1 1456 AA; 163231 MW; 8DE8AAB22BC481P2 CRC64;  
 Query Match 74.1%; Score 7419.5; DB 2; Length 1456;  
 Best Local Similarity 96.1%; Pred. No. 0;  
 Matches 1456; Conservative 0; Mismatches 0; Indels 59; Gaps 1;  
 QY 17 DNNVTNTSNDPFTKCFONTLVWPCFYLWACFPFYLYLSRHRDGYIQTPLNKTCTA 76  
 DB 1 DNNVTNTSNDPFTKCFONTLVWPCFYLWACFPFYLYLSRHRDGYIQTPLNKTCTA 60  
 QY 77 LGFLLIWVCWADLFYSFWSERSGIFLAPVFLVSPTLGLITTLATFLIQLERRKGVQSSG 136  
 DB 61 LGFLLIWVCWADLFYSFWSERSGIFLAPVFLVSPTLGLITTLATFLIQLERRKGVQSSG 120  
 QY 137 IMLTFWLVALCALAILRSKINTALKEDAQVDLFRDITFYVYFSLILLIQLVLSCSFDRSP 196  
 DB 121 IMLTFWLVALCALAILRSKINTALKEDAQVDLFRDITFYVYFSLILLIQLVLSCSFDRSP 180  
 QY 197 LFSETHDNPCESSASPLSRITFWITGLIVRGVROPLEGSDLSLWLNKEDTSQVVPV 256  
 DB 181 LFSETHDNPCESSASPLSRITFWITGLIVRGVROPLEGSDLSLWLNKEDTSQVVPV 240  
 QY 257 LVKNKKCEAKTRKQPKVYVYSSKDPAPKSSSKVDANEEBALIVKSPQKEMNPSLFKV 316  
 DB 241 LVKNKKCEAKTRKQPKVYVYSSKDPAPKSSSKVDANEEBALIVKSPQKEMNPSLFKV 300  
 QY 317 LYKTPGYPFLMSEFFKAHDLMMFGSPQILKLLIKFVNDTKAPDWQGYFTVLLFTACL 376  
 DB 301 LYKTPGYPFLMSEFFKAHDLMMFGSPQILKLLIKFVNDTKAPDWQGYFTVLLFTACL 360  
 QY 377 QTLVLHGVPHICFVSGMRITKAVIGAVYRKALVITNSARKSTVGEIYNLMSVDQAFMD 436  
 DB 361 QTLVLHGVPHICFVSGMRITKAVIGAVYRKALVITNSARKSTVGEIYNLMSVDQAFMD 420  
 QY 437 LATYINMIWSAPLOVILALYLWNLGPSVLGAVAVMLVMPVNAVMAKTKTYQVAHMK 496  
 DB 421 LATYINMIWSAPLOVILALYLWNLGPSVLGAVAVMLVMPVNAVMAKTKTYQVAHMK 480  
 QY 497 SKDNRIKLMNEILNGIKVLKAWELA FKQKVLATROBELKVLKKSAYLSAVGTFTTWCT 556  
 DB 481 SKDNRIKLMNEILNGIKVLKAWELA FKQKVLATROBELKVLKKSAYLSAVGTFTTWCT 540  
 QY 557 PELVALCTFAVVTIDENNILDAQTAFAVSLALFNILRPPLNLPWVISSIVQASVSLKRL 616  
 DB 541 PELVALCTFAVVTIDENNILDAQTAFAVSLALFNILRPPLNLPWVISSIVQASVSLKRL 600  
 QY 617 RIFLSHEELEPDSIERPPVKDGGTNSITVRNATFTWARSDDPTLNGITFTSPEGALNAV 676  
 DB 601 RIFLSHEELEPDSIERPPVKDGGTNSITVRNATFTWARSDDPTLNGITFTSPEGALNAV 660  
 QY 677 VQVGGCKGKSLLSALLAEMDKVEGHVAIKGSVAYVPOQAWIQNDSLRNILEFGCQLEBPY 736  
 DB 661 VQVGGCKGKSLLSALLAEMDKVEGHVAIKGSVAYVPOQAWIQNDSLRNILEFGCQLEBPY 688  
 QY 737 YRSVIQACALLPDLEILPSGDRTEIGEKGWNLGGOKQVRSVLARAVYSNADYLPDDPLS 796  
 DB 689 -----KGWNLGGOKQVRSVLARAVYSNADYLPDDPLS 721  
 QY 797 AVDAHVGKHFENVIGPKGMLKNKTRILVTHSMYLPQVDVITVMSGGKISEMGVQELL 856  
 DB 722 AVDAHVGKHFENVIGPKGMLKNKTRILVTHSMYLPQVDVITVMSGGKISEMGVQELL 781  
 QY 857 ARDGAFAEFLRYASTEQDAENGVTGSGPGKEAKQWENGMLVYTDGAGLQRLQSS 916  
 DB 782 ARDGAFAEFLRYASTEQDAENGVTGSGPGKEAKQWENGMLVYTDGAGLQRLQSS 841

QY 917 SSSYSGDISRHNSHTAELOKAEAKKEETWKLMEADKAOTGOVKLSVYWDYMKAIGLFISF 976  
 DB 842 SSSYSGDISRHNSHTAELOKAEAKKEETWKLMEADKAOTGOVKLSVYWDYMKAIGLFISF 901  
 QY 977 LSIPLFCMNVHSALASANYLWLSLWTDPIVNGTOBHTKVRLSVYGALGTSQGIIVFGYMA 1036  
 DB 902 LSIPLFCMNVHSALASANYLWLSLWTDPIVNGTOBHTKVRLSVYGALGTSQGIIVFGYMA 961  
 QY 1037 VSIIGILASRCLHVDLHLSILRSWSPSPERTPSGNLVNRFSEKELDTVDMSIPEVIMKFMFG 1096  
 DB 962 VSIIGILASRCLHVDLHLSILRSWSPSPERTPSGNLVNRFSEKELDTVDMSIPEVIMKFMFG 1021  
 QY 1097 SLFNVIAGACIVILLATPIAAIIIPPLGLIYFFVORFYVASSRQLKRLSVSRSPVYSHFN 1156  
 DB 1022 SLFNVIAGACIVILLATPIAAIIIPPLGLIYFFVORFYVASSRQLKRLSVSRSPVYSHFN 1081  
 QY 1157 ETLGVSIVIRAFESQERFIHQSDLKVDENQKAYYPSIVANRWLAVRLECVGNCIVLFAAL 1216  
 DB 1082 ETLGVSIVIRAFESQERFIHQSDLKVDENQKAYYPSIVANRWLAVRLECVGNCIVLFAAL 1141  
 QY 1217 FAVISRHSLSAGLGLSVSYSLQVTTYLNWLVRMSSEMETNI VAVERLKEYSETEKEAPW 1276  
 DB 1142 FAVISRHSLSAGLGLSVSYSLQVTTYLNWLVRMSSEMETNI VAVERLKEYSETEKEAPW 1201  
 QY 1277 QIQTAPPSSWPQVGRVEFRNYCLRYREDLDFVLRHINVTINGGKVGIVGRTGAGKSSL 1336  
 DB 1202 QIQTAPPSSWPQVGRVEFRNYCLRYREDLDFVLRHINVTINGGKVGIVGRTGAGKSSL 1261  
 QY 1337 TLGLFRINESAGEHIIIDGINIAKGLHDLRFKITIIPQDPVLPFGSLRMNLDPPSOYSD 1396  
 DB 1262 TLGLFRINESAGEHIIIDGINIAKGLHDLRFKITIIPQDPVLPFGSLRMNLDPPSOYSD 1321  
 QY 1397 BEVWTSLEHLKDFVSALPKDLHDECAEGENLSVGORQLVCLARALLRKTILVLDEA 1456  
 DB 1322 BEVWTSLEHLKDFVSALPKDLHDECAEGENLSVGORQLVCLARALLRKTILVLDEA 1381  
 QY 1457 TAAVLETDLTIQSTIRTFQEDCTVLTIAHRLNTIMDYTRVILVDKGIQIEYGAPSDLLQ 1516  
 DB 1382 TAAVLETDLTIQSTIRTFQEDCTVLTIAHRLNTIMDYTRVILVDKGIQIEYGAPSDLLQ 1441  
 QY 1517 QRGLFYSMAKDAGLV 1531  
 DB 1442 QRGLFYSMAKDAGLV 1456  
 RESULT 7  
 Q6UR05 PRELIMINARY; PRT; 1531 AA.  
 AC Q6UR05;  
 DT 05-JUL-2004 (TREMELrel. 27, Created)  
 DT 05-JUL-2004 (TREMELrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TREMELrel. 27, Last annotation update)  
 DE Multidrug resistance-associated protein 1.  
 GN Name=MRP1;  
 OS Canis familiaris (Dog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 OC NCBI\_TaxId=9615;  
 RN [1]  
 SEQUENCE FROM N.A.  
 RX MEDLINE=22404678; PubMed=12516967;  
 RA Ma L., Pratt S.E., Cao J., Dantzig A.H., Moore R.E., Slapak C.A.;  
 RT "Identification and characterization of the canine multidrug  
 RT resistance-associated protein";  
 RL Mol. Cancer Ther. 1:1335-1342(2002).  
 RN [2]  
 SEQUENCE FROM N.A.  
 RA Ma L., Pratt S.E., Cao J., Dantzig A.H., Moore R.E., Slapak C.A.;  
 RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: Belongs to the ABC transporter family.  
 DR EMBL; AY363728; AAQ3148.1; --  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0005524; F:ATP binding; IEA.

DR GO: 0042626; F:ATPase activity, coupled to transmembrane m. . . ; IEA.  
 DR GO: 0000166; F:nucleotide binding; IEA.  
 DR GO: 0005215; F:transporter activity; IEA.  
 DR GO: 0006810; F:transport; IEA.  
 DR InterPro: IPR003593; AAA ATPase.  
 DR InterPro: IPR011527; ABC membrane 1.  
 DR InterPro: IPR001140; ABC\_TM\_transp.  
 DR InterPro: IPR003439; ABC\_transporter.  
 DR InterPro: IPR001395; Aldo/ket\_red.  
 DR InterPro: IPR005292; MRP\_assoc.  
 DR InterPro: IPR000719; Prot\_kinase.  
 DR Pfam: PF00664; ABC membrane; 2.  
 DR Pfam: PF00005; ABC trans; 2.  
 DR ProDom: PD000006; ABC transporter; 2.  
 DR SMART: SM00382; AAA; 2.  
 DR TIGRfam: TIGR00957; MRP assoc\_pro; 1.  
 DR PROSITE: PS00929; ABC TMIF; 2.  
 DR PROSITE: PS00211; ABC\_TRANSPORTER\_1; 2.  
 DR PROSITE: PS00893; ABC\_TRANSPORTER\_2; 2.  
 DR PROSITE: PS00063; ALDO-KETO REDUCTASE\_3; UNKNOWN\_1.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; UNKNOWN\_1.  
 KW ATP-binding.  
 SQ SEQUENCE 1531 AA; 171790 MW; 746361A71C6158BD CRC64;  
 Query Match 72.6%; Score 7272; DB 2; Length 1531;  
 Best Local Similarity 91.6%; Pred. No. 0;  
 Matches 1402; Conservative 70; Mismatches 59; Indels 0; Gaps 0;  
 QY 1 MALRGFCGADGSDPDLWNNVTWNTSNPDFTKCFQNTVLVWPCFYLWACFPFFYFLYLSRH 60  
 DB 1 MALRGFCGADGSDPDLWNNVTWNTSNPDFTKCFQNTVLVWPCFYLWACFPFFYFLYLSRH 60  
 QY 61 DRYGTQMTPLNKTALGFLWIVCWADLFYFWSRSGIFLAPVPLVSPILLGITLLA 120  
 DB 61 DRYGTQMTPLNKTALGFLWIVCWADLFYFWSRSGIFLAPVPLVSPILLGITLLA 120  
 QY 121 TFLIQLERRKGQSSGIMLTFWLVALCALATLRSKIMTALKEDAQVDFLDITFYVYFS 180  
 DB 121 TFLIQLERRKGQSSGIMLTFWLVALCALATLRSKIMTALKEDAQVDFLDITFYVYFS 180  
 QY 181 LLLIQLVLSFSDRPLSETHDNPCESSASFLSRITFWITGLIVRGYRPLEGSD 240  
 DB 181 LLLIQLVLSFSDRPLSETHDNPCESSASFLSRITFWITGLIVRGYRPLEGSD 240  
 QY 241 LWSLNKEDTSEQVPLVKNMKCECAKTRQPKVYSSKDPAPKESKVDANEEVAL 300  
 DB 241 LWSLNKEDTSEQVPLVKNMKCECAKTRQPKVYSSKDPAPKESKVDANEEVAL 300  
 QY 301 IVKSPQKEWNPFLFKVLYKTGYPFLMSFFFKAIHDLMMFGSPQITLKLIFVNDTKAPD 360  
 DB 301 IVKTPQKEREPSLFLVLYKTGYPFLMSFLFKALHDLMMFAGPELKLILINFNVDKKAPD 360  
 QY 361 WQGYPTVLLFVTAQLVLHQYPHICFVSGMRITKAVIGAVYRKALVITNSARKSTV 420  
 DB 361 WQGYLTALLFTACQLVLHQYPHICFVSGMRITKAVIGAVYRKALVITNSARKSTV 420  
 QY 421 GEIVNLSVDAQRFMDLATYINMTWSAPLOVILALYLWNLGSPVLGAVVWMLVWPN 480  
 DB 421 GEIVNLSVDAQRFMDLATYINMTWSAPLOVILALYLWNLGSPVLGAVVWMLVWPN 480  
 QY 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGI KVLKLYAWELAFKDKVLAIROBELKVLK 540  
 DB 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGI KVLKLYAWELAFKDKVLAIROBELKVLK 540  
 QY 541 KSAYLSAVGTTWCTPFLVALCTFAVVTVDENNILDAQAFVSLAFNLIRPFLNLP 600  
 DB 541 KSAYLSAVGTTWCTPFLVALCTFAVVTVDENNILDAQAFVSLAFNLIRPFLNLP 600  
 QY 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIRRRPVKDGSGTNSITVRNATFTWARSOPPT 660  
 DB 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIRRRPVKDGSGTNSITVRNATFTWARSOPPT 660  
 QY 661 LINGITFSIPEGALVAVVGQVCGKSSLLSALLAEMDKVEGHVAIKGSVAYVPPQAWIQND 720

DB 661 LSGITFSIPEGALVAVVGQVCGKSSLLSALLAEMDKVEGHVAIKGSVAYVPPQAWIQND 720  
 QY 721 SLRENILFGCOLLEPYRSVIAQCALLPDLBELPSGDRTEIGEGKGNLSGQKQKRVSLAR 780  
 DB 721 SLRENILFGCOLLEPYRSVIAQCALLPDLBELPSGDRTEIGEGKGNLSGQKQKRVSLAR 780  
 QY 781 AVYNSADIYLPDDPLSAVDHVGKHI FENVIGPKGMLKNKTRILVTHSMYSYLPQVDVIV 840  
 DB 781 AVYNSADIYLPDDPLSAVDHVGKHI FENVIGPKGMLKNKTRILVTHSMYSYLPQVDVIV 840  
 QY 841 MSGGKISEMGSYQELLARDGAFELRTYASTEOEADAENGVTGCVSGPGKEAKOMENG 900  
 DB 841 MSGGKISEMGSYQELLARDGAFELRTYASTEOEADAENGVTGCVSGPGKEAKOMENG 900  
 QY 901 LVTDSAGQLQRLSSSSSYSGDISRHHNSTAEILQKAEAKKEETWKLMEADKAQTQGVKL 960  
 DB 901 LVTDSAGQLQRLSSSSSYSGDISRHHNSTAEILQKAEAKKEETWKLMEADKAQTQGVKL 960  
 QY 961 SVYDYMKAIGLIFSLIFLFCNHNVSALASNTWLSLWTDPIVNGTQEHKTVRLSVYG 1020  
 DB 961 SVYDYMKAIGLIFSLIFLFCNHNVSALASNTWLSLWTDPIVNGTQEHKTVRLSVYG 1020  
 QY 1021 ALGISQGIATVFGYSMAVSIIGGILASRCLHVDLLHSILRSFMSFFERTPSGNLVNRESKEL 1080  
 DB 1021 ALGISQGIATVFGYSMAVSIIGGILASRCLHVDLLHSILRSFMSFFERTPSGNLVNRESKEL 1080  
 QY 1081 DTVDSMIPKMFMSGLFNIVIGACIVILLATPIAAIIIPPLGLIYFFVQFYVASSRQL 1140  
 DB 1081 DTVDSMIPKMFMSGLFNIVIGACIVILLATPIAAIIIPPLGLIYFFVQFYVASSRQL 1140  
 QY 1141 KRLESVSRSPYSHFNELTLLGVSVIRAFESQERFIRQSLKVDENOKAYTSPISVANRWLA 1200  
 DB 1141 KRLESVSRSPYSHFNELTLLGVSVIRAFESQERFIRQSLKVDENOKAYTSPISVANRWLA 1200  
 QY 1201 VRLECVGNCIVLFAALFAVISRHSLSAGLSYSLSQVTTYNLWLRMSSEMETNIVA 1260  
 DB 1201 VRLECVGNCIVLFAALFAVISRHSLSAGLSYSLSQVTTYNLWLRMSSEMETNIVA 1260  
 QY 1261 VERLKEYSETEKEAPWQIQTAPPSSWPQVGRVFRNYCLRYREDLDFVLRHINVTINGG 1320  
 DB 1261 VERLKEYSETEKEAPWQIQTAPPSSWPQVGRVFRNYCLRYREDLDFVLRHINVTINGG 1320  
 QY 1321 EKVICVGTGAGKSSLTGLFRINESAEGEIIIDGINAKIGLHDLAFKTIIPQDPVLF 1380  
 DB 1321 EKVICVGTGAGKSSLTGLFRINESAEGEIIIDGINAKIGLHDLAFKTIIPQDPVLF 1380  
 QY 1381 SGSLRMNLDPPSOYSDSEVWTSLELAHLKDFVSALPKDLHCECAEGENLSVGQRLVCL 1440  
 DB 1381 SGSLRMNLDPPSOYSDSEVWTSLELAHLKDFVSALPKDLHCECAEGENLSVGQRLVCL 1440  
 QY 1441 ARALLRKTILVDEATAAVALDETDLLIQSTIRTFQEDCTVLTIAHRLNTIMDYTRIVL 1500  
 DB 1441 ARALLRKTILVDEATAAVALDETDLLIQSTIRTFQEDCTVLTIAHRLNTIMDYTRIVL 1500  
 QY 1501 DKGEIQSVGAPSLDLOQRLGYFSMAKDAGLV 1531  
 DB 1501 DKGEIQSVGAPSLDLOQRLGYFSMAKDAGLV 1531  
 RESULT 8  
 Q8HXQ5  
 ID Q8HXQ5 PRELIMINARY; PRT; 1530 AA.  
 AC Q8HXQ5;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Multidrug resistance protein 1.  
 GN Name=MRP1;  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovinae; Bos.

NCBI\_TaxID=9913;  
 (1)  
 SEQUENCE FROM N.A.  
 MEDLINE=22062955; PubMed=12067707; DOI=10.1016/S0014-5793(02)02816-8;  
 RA Taguchi Y., Saeki K., Komano T.;  
 RT "Functional analysis of MRPI cloned from bovine."  
 RL FEBS Lett. 521:211-213(2002).  
 CC -!- SIMILARITY: Belongs to the ABC transporter family.  
 DR EMBL: AB082124; BAC15550.1; -;  
 DR HSSP: P08716; 1MT0.  
 DR GO: GO:0016021; C: integral to membrane; IEA.  
 DR GO: GO:0005524; F: ATP binding; IEA.  
 DR GO: GO:0042626; F: ATPase activity, coupled to transmembrane m. . .; IEA.  
 DR GO: GO:000166; F: nucleotide binding; IEA.  
 DR GO: GO:0005215; F: transporter activity; IEA.  
 DR GO: GO:0006810; P: transport; IEA.  
 DR InterPro: IPR003593; AAA ATPase.  
 DR InterPro: IPR011527; ABC membrane 1.  
 DR InterPro: IPR001140; ABC TM transp.  
 DR InterPro: IPR003439; ABC transporter.  
 DR InterPro: IPR005292; MRP\_assoc.  
 DR Pfam: PF00664; ABC membrane; 2.  
 DR Pfam: PF00005; ABC\_tran; 2.  
 DR ProDom: PD000006; ABC transporter; 2.  
 DR SMART: SM00382; AAA; 2.  
 DR TIGRFAMs: TIGR00957; MRP assoc\_pro; 1.  
 DR PROSITE: PS00929; ABC TMIF; 2.  
 DR PROSITE: PS00211; ABC\_TRANSPORTER\_1; 2.  
 DR PROSITE: PS00893; ABC\_TRANSPORTER\_2; 2.  
 KW ATP-binding.  
 SQ SEQUENCE 1530 AA; 171664 MW; AAE4F92ED7832703 CRC64;  
 Query Match 72.2%; Score 7230.5; DB 2; Length 1530;  
 Best Local Similarity 90.7%; Pred No. 0;  
 Matches 1388; Conservative 83; Mismatches 59; Indels 1; Gaps 1;  
 QY 1 MALRGCSDGSDPLMDNWNVTNTPDFTKCFQNTLVVWPCFYLVACPPFYLYLSRH 60  
 DB 1 MALRDFCSVDGSDLPFWNNVTNTPDFTKCFQNTLVVWPCFYLVACPPFYLYLSRH 60  
 QY 61 DRGYIQTPLNKTALGFLWIVCWADLFVSWERSGIFLAPVFLSPILLGTTLLA 120  
 DB 61 DRGYIQTPLNKTALGFLWIVCWADLFVSWERSGIFLAPVFLSPILLGTTLLA 120  
 QY 121 TFLIQLERRKGQSSGIMLTFWLVALCALATLRSKIMTALKEDAQVDFRDIYVYVFS 180  
 DB 121 TFLIQLERRKGQSSGIMLTFWLVALCALATLRSKIMTALKEDAQVDFRDIYVYVFS 180  
 QY 181 LLLIQLVLSGSDRSLFSETHDPNCPPESSASFLSRITFWITGLIVRGYRQPLEGSD 240  
 DB 181 LLLIQLVLSGSDRSLFSETHDPNCPPESSASFLSRITFWITGLIVRGYRQPLEGSD 240  
 QY 241 LWSLNKEDTSEQVPLVKNWKECAKTRKQPVVYSSKDPAPKESKVDANEVEAL 300  
 DB 241 LWSLNKEDTSEQVPLVKNWKECAKTRKQPVVYSSKDPAPKESKVDANEVEAL 300  
 QY 301 IVKSPQKEWNSLFVLYKTPGYPFLMFPFKAHDLAMFSGPQLTKLLIKFVNDTKAPD 360  
 DB 301 IVKSPQKEWNSLFVLYKTPGYPFLMFPFKAHDLAMFSGPQLTKLLIKFVNDTKAPD 360  
 QY 361 WQGYFYTVLLFVTACIQLTLVHLQYFHI CFVSGMRKTAIVAGYRKALVITNSARKSTV 420  
 DB 361 WQGYFYTVLLFVTACIQLTLVHLQYFHI CFVSGMRKTAIVAGYRKALVITNSARKSTV 420  
 QY 421 GEIVNLSVDAQRFDLATYINNIWSAPLQVTLALYLLWNLGSPVLGAVVWMLVMPVN 480  
 DB 421 GEIVNLSVDAQRFDLATYINNIWSAPLQVTLALYLLWNLGSPVLGAVVWMLVMPVN 480  
 QY 481 AVAMKTKTYQVAHWKSKDNRIKLWNEILNGIKVLKYAWELAFKDKVLAIROBELKVLK 540  
 DB 481 AVAMKTKTYQVAHWKSKDNRIKLWNEILNGIKVLKYAWELAFKDKVLAIROBELKVLK 540  
 QY 541 KSAYLSAVGTFTVWCTPFLVALCTFAVYVTTDENNILDQAFAVSLALFNILRFPFLNLP 600

DB 541 KSAYLSAVGTFTVWCTPFLVALCTFAVYVTTDENNILDQAFAVSLALFNILRFPFLNLP 600  
 QY 601 MVISSIVQASVSLKRLRIFLSHEELEPDSEIRRPVKDGGGTTNSITVRATFTWARSDDPT 660  
 DB 601 MVISSIVQASVSLKRLRIFLSHEELEPDSEIRRPVKDGGGTTNSITVRATFTWARSDDPT 660  
 QY 661 LINGITSIPEGALVAVVQVCGCGKSSLLSALLAEMDKVEGHVIAIKGSVAYVPPQAWIQND 720  
 DB 661 LINGITSIPEGALVAVVQVCGCGKSSLLSALLAEMDKVEGHVIAIKGSVAYVPPQAWIQND 720  
 QY 721 SLRENILFGCOLEPEYRYSVIOACALLDILPSGDRTEIGEKVNLSSGQKQVSLAR 780  
 DB 721 SLRENILFGCOLEPEYRYSVIOACALLDILPSGDRTEIGEKVNLSSGQKQVSLAR 780  
 QY 781 AVYSNADIYLPDDPLSAVDHVGKHPENVIIPGKMLKNTKRLIVTHSMVSLPQVDVIV 840  
 DB 781 AVYSNADIYLPDDPLSAVDHVGKHPENVIIPGKMLKNTKRLIVTHSMVSLPQVDVIV 840  
 QY 841 MSGGKISEMGSYQELLARDGAFELRTYASTEQEQDAEENGVTGVSFGPKGKAKOMENG 900  
 DB 841 MSGGKISEMGSYQELLARDGAFELRTYASTEQEQDAEENGVTGVSFGPKGKAKOMENG 900  
 QY 901 LVTSAGKQLOQLSSSSSYSGDISRHNSHNSAEKAEKKEETWKLMEADKAQGVKL 960  
 DB 901 LVTSAGKQLOQLSSSSSYSGDISRHNSHNSAEKAEKKEETWKLMEADKAQGVKL 960  
 QY 961 SVYDYMKAIGLIFSLFISFLFCNHNVSALASNYWLSLWTDPIVNGTQEHKTVLSVYG 1020  
 DB 961 SVYDYMKAIGLIFSLFISFLFCNHNVSALASNYWLSLWTDPIVNGTQEHKTVLSVYG 1020  
 QY 1021 ALGISQGIAGVFGYSMAVSGIGLILSRCLVHLLHSILRSPMSFFERTPSGNLVNRFSEL 1080  
 DB 1021 ALGISQGIAGVFGYSMAVSGIGLILSRCLVHLLHSILRSPMSFFERTPSGNLVNRFSEL 1080  
 QY 1081 DTVDSPMPEVTKMFGSLFNIVIGACIVILLATPAAIIIPGLGLIYFFVQRYVASSRQL 1140  
 DB 1081 DTVDSPMPEVTKMFGSLFNIVIGACIVILLATPAAIIIPGLGLIYFFVQRYVASSRQL 1140  
 QY 1141 KRLSVSRSPVYSHFNETHLLGVSVTRAFEEQERFIHQSDLVKVDENQKAYYPSIVANRWLA 1200  
 DB 1141 KRLSVSRSPVYSHFNETHLLGVSVTRAFEEQERFIHQSDLVKVDENQKAYYPSIVANRWLA 1200  
 QY 1201 VRLCEVGNICVLFALFAVI SRHSLSAGLVSLYSYSLQVTTYLNLVMSSEMETNI VA 1260  
 DB 1201 VRLCEVGNICVLFALFAVI SRHSLSAGLVSLYSYSLQVTTYLNLVMSSEMETNI VA 1260  
 QY 1261 VERLKEYSETEKEAPWQIQTAPPSSWQVGRVFRNCLYREDLDFVLRHINTINGG 1320  
 DB 1261 VERLKEYSETEKEAPWQIQTAPPSSWQVGRVFRNCLYREDLDFVLRHINTINGG 1320  
 QY 1321 EKVGVIGRTGAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
 DB 1321 EKVGVIGRTGAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
 QY 1381 SGSLRMNLDPSQYSDSEEVWTSLELAHLKDFVSALPDKLDHECAGGENLSVQRLVCL 1440  
 DB 1381 SGSLRMNLDPSQYSDSEEVWTSLELAHLKDFVSALPDKLDHECAGGENLSVQRLVCL 1440  
 QY 1441 ARALLRKTKIILVDSATAVDETDLLIQSTIRTOFEDCTVTIAHRLNTIMDYTRVIVL 1500  
 DB 1441 ARALLRKTKIILVDSATAVDETDLLIQSTIRTOFEDCTVTIAHRLNTIMDYTRVIVL 1500  
 QY 1501 DKGEIQEWGAPSDLLQQRGLFYSMADKAGLV 1531  
 DB 1501 DKGEIQEWGAPSDLLQQRGLFYSMADKAGLV 1531

RESULT 9  
 Q9UQ98 PRELIMINARY; PRT; 1400 AA.  
 AC Q9UQ98;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)



QY 1157 ETLLGVSVIRAFEEQERITHOSDLKVDENQKAYPSIVANRWLAVRLCVGNCIVLFAAL 1216  
 DB 1026 ETLLGVSVIRAFEEQERITHOSDLKVDENQKAYPSIVANRWLAVRLCVGNCIVLFAAL 1085  
 QY 1217 FAVISRHSLSAGLVGLSVYSYLSQVTTYNLWLRMSSEMETNIVAVRLKEYSETEKEAPW 1276  
 DB 1086 FAVISRHSLSAGLVGLSVYSYLSQVTTYNLWLRMSSEMETNIVAVRLKEYSETEKEAPW 1145  
 QY 1277 QIQTAPSSPQVGRVFRNVCRLRYREDLDPVLRHINVTNGGKVGIVGRTGAKGSSL 1336  
 DB 1146 QIQTAPSSPQVGRVFRNVCRLRYREDLDPVLRHINVTNGGKVGIVGRTGAKGSSL 1205  
 QY 1337 TIGLFRINESAGEEIIIDGINIAKIGLHDLRFKTIIPQDPVLFSGSLRMNLDPPFSQYSD 1396  
 DB 1206 TIGLFRINESAGEEIIIDGINIAKIGLHDLRFKTIIPQDPVLFSGSLRMNLDPPFSQYSD 1265  
 QY 1397 BEVWTSLELAHLKOFVSALPKDLHCECAEGGENLSVGQRLVCLARALLRTKILVLDEA 1456  
 DB 1266 BEVWTSLELAHLKOFVSALPKDLHCECAEGGENLSVGQRLVCLARALLRTKILVLDEA 1325  
 QY 1457 TAAVDETDLLIOSTIRTFQEDCTVLTAAHLNTIMDYTRVILDKGIEQVGAPSDLLQ 1516  
 DB 1326 TAAVDETDLLIOSTIRTFQEDCTVLTAAHLNTIMDYTRVILDKGIEQVGAPSDLLQ 1385  
 QY 1517 ORGLFYSMKADAGLV 1531  
 DB 1386 ORGLFYSMKADAGLV 1400

RESULT 10  
 MRP1 MOUSE  
 ID MRP1 MOUSE STANDARD; PRT; 1528 AA.  
 AC O35379;  
 DT 25-OCT-2004 (Rel. 45, Created)  
 DT 25-OCT-2004 (Rel. 45, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Multidrug resistance-associated protein 1 (ATP-binding cassette, sub-family C, member 1).  
 DE family C, member 1).  
 GN Name=Abcc1; Synonyms=Abcc1a, Abcc1b, Mdrap, Mrp;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]\_TaxID=10090;  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Muscle;  
 RX MEDLINE=96251691; PubMed=8649356;  
 RA Stride B.D., Valdimarsson G., Gerlach J.H., Wilson G.M., Cole S.P.,  
 RA Deeley R.G.;  
 RT "Structure and expression of the messenger RNA encoding the murine  
 RT multidrug resistance protein, an ATP-binding cassette transporter.";  
 RL Mol. Pharmacol. 49:962-971(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Testis;  
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;  
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
 RA Nikaide I., Oeato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,  
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,  
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,  
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,  
 RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,  
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,  
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,  
 RA Grimmond S., Gusicinch S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
 RA Kana A., Kawaji H., Kawasaki Y., Kedzierski R.M., King B.L.,  
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,  
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,  
 RA Nagashima T., Nunata K., Okido T., Pavan W.J., Perce G., Pesole G.,  
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,  
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,  
 RA Sandelin A., Schneider C., Semp C.A., Setou M., Shimada K.,

RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,  
 RA Verardo R., Wagner L., Wanlestedt C., Wang Y., Watanabe Y., Wells C.,  
 RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,  
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,  
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
 RA Hara A., Hashizume W., Imotani K., Iehii Y., Itoh M., Kagawa I.,  
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
 RA Birney E., Havaehizaki Y.;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs.";  
 RL Nature 420:563-573(2002).  
 CC -!- FUNCTION: May participate directly in the active transport of  
 CC drugs into subcellular organelles or influence drug distribution  
 CC indirectly (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).  
 CC -!- SIMILARITY: Belongs to the ABC transporter family. MRP subfamily.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AF022908; AAB80938.1; --  
 CC EMBL; AK029896; BAC26654.1; --  
 CC HSSP; P08716; IMT0.  
 CC MGI; MGI:102676; Abcc1.  
 CC InterPro; IPR003593; AAA\_ATPase.  
 CC InterPro; IPR011527; ABC\_membrane\_1.  
 CC InterPro; IPR001140; ABC\_TM\_transp.  
 CC InterPro; IPR003439; ABC\_transporter.  
 CC InterPro; IPR005292; MRP\_assoc.  
 CC Pfam; PF00664; ABC\_membrane; 2.  
 CC Pfam; PF00005; ABC\_tran; 2.  
 CC ProDom; PD000006; ABC\_transporter; 2.  
 CC SMART; SM00382; AAA; 2.  
 CC TIGRFAMs; TIGR00957; MRP\_assoc\_pro; 1.  
 CC PROSITE; PS00929; ABC\_TMIF; 2.  
 CC PROSITE; PS0211; ABC\_TRANSPORTER\_1; 2.  
 CC PROSITE; PS00893; ABC\_TRANSPORTER\_2; 2.  
 KW ATP-binding; Glycoprotein; Repeat; Transmembrane; Transport.  
 FT DOMAIN 1 33 Extracellular (By similarity).  
 FT TRANSMEM 34 54 1 (By similarity).  
 FT DOMAIN 55 74 Cytoplasmic (By similarity).  
 FT TRANSMEM 75 95 2 (By similarity).  
 FT DOMAIN 96 100 Extracellular (By similarity).  
 FT TRANSMEM 101 121 3 (By similarity).  
 FT DOMAIN 122 133 Cytoplasmic (By similarity).  
 FT TRANSMEM 134 154 4 (By similarity).  
 FT DOMAIN 155 172 Extracellular (By similarity).  
 FT TRANSMEM 173 193 5 (By similarity).  
 FT DOMAIN 194 317 Cytoplasmic (By similarity).  
 FT TRANSMEM 318 338 6 (By similarity).  
 FT DOMAIN 339 364 Extracellular (By similarity).  
 FT TRANSMEM 365 385 7 (By similarity).  
 FT DOMAIN 386 441 Cytoplasmic (By similarity).  
 FT TRANSMEM 442 462 8 (By similarity).  
 FT DOMAIN 463 486 Extracellular (By similarity).  
 FT TRANSMEM 487 548 9 (By similarity).  
 FT DOMAIN 549 569 10 (By similarity).  
 FT TRANSMEM 570 591 Extracellular (By similarity).  
 FT TRANSMEM 592 612 11 (By similarity).  
 FT DOMAIN 613 963 Cytoplasmic (By similarity).  
 FT TRANSMEM 964 984 12 (By similarity).  
 FT DOMAIN 985 1022 Extracellular (By similarity).  
 FT TRANSMEM 1023 1043 13 (By similarity).  
 FT DOMAIN 1044 1086 Cytoplasmic (By similarity).  
 FT TRANSMEM 1087 1107 14 (By similarity).

```

FT DOMAIN 1108 1108 Extracellular (By similarity).
FT TRANSMEM 1109 1129 15 (By similarity).
FT DOMAIN 1130 1200 Cytoplasmic (By similarity).
FT TRANSMEM 1201 1220 16 (By similarity).
FT DOMAIN 1221 1233 Extracellular (By similarity).
FT TRANSMEM 1234 1244 17 (By similarity).
FT DOMAIN 1245 1528 Cytoplasmic (By similarity).
FT TRANSMEM 1246 1528 ABC transporter 1.
FT DOMAIN 1247 1528 ABC transporter 2.
FT NP_BIND 1248 1528 ATP (potential).
FT NP_BIND 1331 1331 ATP (potential).
FT CARBOHYD 19 19 N-linked (GlcNAc...) (potential).
FT CARBOHYD 1003 1003 N-linked (GlcNAc...) (potential).
SQ SEQUENCE 1528 AA; 171183 MW; 68FD13667D61DBBB CRC64;

Query Match 69.9%; Score 7002.5; DB 1; Length 1528;
Best Local Similarity 88.0%; Pred. No. 0;
Matches 1349; Conservative 102; Mismatches 75; Indels 7; Gaps 5;

QY 1 MALRGCAGDGDPLDNDNVTWNTSNDPFTKCFQNTLVWVPCFYLWACFPFYFLYLSRH 60
DB 1 MALRGCAGDGDPLDNDNVTWNTSNDPFTKCFQNTLVWVPCFYLWACFPFYFLYLSRH 60
QY 61 DRGYIQMTPLNKTALGFLWIVCWADLFYSFWERSRGIFLAPVFLVSPITLLGITLLA 120
DB 61 DRGYIQMTPLNKTALGFLWIVCWADLFYSFWERSRGIFLAPVFLVSPITLLGITLLA 120
QY 121 TFLIOLERRKGVOSSGIMTLFWALVALCALAILRSKIMTALKEDAQVDLFRDITPVYVFS 180
DB 121 TFLIOLERRKGVOSSGIMTLFWALVALCALAILRSKIMTALKEDAQVDLFRDITPVYVFS 180
QY 181 LLLIQLVLSCFSDRPLSETHDNPCESSASFLSRITFWITGLVIRGVROPLEGSD 240
DB 181 LLLIQLVLSCFSDRPLSETHDNPCESSASFLSRITFWITGLVIRGVROPLEGSD 240
QY 241 LWSLNKEDTSEQVPELVLYNKKCEKATRKOPKVVYVS - SKDPAQKESKSKVDANEVEEA 299
DB 241 LWSLNKEDTSEQVPELVLYNKKCEKATRKOPKVVYVS - SKDPAQKESKSKVDANEVEEA 299
QY 300 LIVKSPQKWNPSLFVLYKTFGPFYFLMSFFFKAHDLMMFSGPOILKLLIKFVNDTKAP 359
DB 301 LIVKSPQKWNPSLFVLYKTFGPFYFLMSFFFKAHDLMMFSGPOILKLLIKFVNDTKAP 359
QY 360 DWQGYFYTVLLFVTAQTLVLHVFHICFVSGMRKTAIVGAVYRKALVITNSARKSST 419
DB 361 DWQGYFYTVLLFVTAQTLVLHVFHICFVSGMRKTAIVGAVYRKALVITNSARKSST 419
QY 420 VGEIVNLSVDAQRPMDLATYINNIWSAPLOVILALYLLNLGSPVLAVVAVMLVMPV 479
DB 421 VGEIVNLSVDAQRPMDLATYINNIWSAPLOVILALYLLNLGSPVLAVVAVMLVMPV 479
QY 480 NAVVAMTKTKYQVAHMKSKDNRIKLMNEILNGIKVLYAYELAFKDKVLAIRQELKVL 539
DB 481 NAVVAMTKTKYQVAHMKSKDNRIKLMNEILNGIKVLYAYELAFKDKVLAIRQELKVL 539
QY 540 KKSAYLSAVGFTTWCPTPLVALCTFAVYVITDENNILDATAFVSLALFNILAFPLNIL 599
DB 541 KKSAYLSAVGFTTWCPTPLVALCTFAVYVITDENNILDATAFVSLALFNILAFPLNIL 599
QY 600 PMWISSIVQASVSLKRLRIFLSHELEPDSTERRPVKDGCGTNSITVRNATFTTWARSDDP 659
DB 601 PMWISSIVQASVSLKRLRIFLSHELEPDSTERRPVKDGCGTNSITVRNATFTTWARSDDP 659
QY 660 TLNGITFSPREGALVAVVGQVCGCKSLLSALLAEMDKVEGHVAIKGSVAVYVQQAION 719
DB 660 TLNGITFSPREGALVAVVGQVCGCKSLLSALLAEMDKVEGHVAIKGSVAVYVQQAION 719
QY 720 DSLRENILFGQLEBPYRSVIOACALLPDLLEILPSGDRTEIGKGNVLSGGQKQVSLA 779
DB 720 DSLRENILFGQLEBPYRSVIOACALLPDLLEILPSGDRTEIGKGNVLSGGQKQVSLA 779
QY 780 RAVYSNADIYLFDDPLSAVDHVGKHIFENVIGPKGMLKNKTRILVTHSMZYLPQVDVII 839

```

```

DB 780 RAVYSNADIYLFDDPLSAVDHVGKHIFENVIGPKGMLKNKTRILVTHSMZYLPQVDVII 839
QY 840 VMSGKIKISEMGSYQELLARDGAFABFLRTYASTEQEODAEENGVTGVSPPGKEAKOMENG 899
DB 840 VMSGKIKISEMGSYQELLARDGAFABFLRTYASTEQEODAEENGVTGVSPPGKEAKOMENG 899
QY 900 MLVTDTSAGKQORQLSSSSSYSGDISRHHNSTAELOKAEKEETWKLMEADKATQGVK 959
DB 900 MLVTDTSAGKQORQLSSSSSYSGDISRHHNSTAELOKAEKEETWKLMEADKATQGVK 959
QY 960 LSVYDYMKAICGLFISFLSIFLFCMCHVSALASNYWLSLWTD - PIVNGTQSHTKVRLSV 1018
DB 960 LSVYDYMKAICGLFISFLSIFLFCMCHVSALASNYWLSLWTD - PIVNGTQSHTKVRLSV 1018
QY 996 LSVYDYMKAICGLFISFLSIFLFCMCHVSALASNYWLSLWTD - PIVNGTQSHTKVRLSV 1018
DB 996 LSVYDYMKAICGLFISFLSIFLFCMCHVSALASNYWLSLWTD - PIVNGTQSHTKVRLSV 1018
QY 1019 YGALGISQGIIVFGYSMAVSIIGGILASCLHVDLLHSILRSPMSFFERTPGNLVNRFSK 1078
DB 1019 YGALGISQGIIVFGYSMAVSIIGGILASCLHVDLLHSILRSPMSFFERTPGNLVNRFSK 1078
QY 1079 ELDTVDSDMIPVIVKMFWSLNFVIGACIVILLATPIAAIIIPPLGLIYFFVQRFVASSR 1138
DB 1079 ELDTVDSDMIPVIVKMFWSLNFVIGACIVILLATPIAAIIIPPLGLIYFFVQRFVASSR 1138
QY 1139 QLKRLSVSRSPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENOKAYVPSIVANRW 1198
DB 1139 QLKRLSVSRSPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENOKAYVPSIVANRW 1198
QY 1199 LAVRLECVGNCIVLFAALFAVISHSLSAGLVGLSVSYSLQVTVTYLNLVNRMSSEMTNI 1258
DB 1199 LAVRLECVGNCIVLFAALFAVISHSLSAGLVGLSVSYSLQVTVTYLNLVNRMSSEMTNI 1258
QY 1259 VAVERLKEYSTETEAPWQIOETAPPSPQVQVGFERNYCLRYREDLDFVLRLHNTVIN 1318
DB 1259 VAVERLKEYSTETEAPWQIOETAPPSPQVQVGFERNYCLRYREDLDFVLRLHNTVIN 1318
QY 1319 GGEKGVIGRTGAGKSSITLGLFRINSAEGLIIGDINIAGIHLDRFKITTIIPQDPV 1378
DB 1319 GGEKGVIGRTGAGKSSITLGLFRINSAEGLIIGDINIAGIHLDRFKITTIIPQDPV 1378
QY 1379 LFSGLSLRNLDPPFQYSDSEEVWTSLEHLKDFVSALPDKLDHCEAGENLSVGQRLV 1438
DB 1379 LFSGLSLRNLDPPFQYSDSEEVWTSLEHLKDFVSALPDKLDHCEAGENLSVGQRLV 1438
QY 1439 CLARALLRKTKILVLDATAVLETDLDLQISTRTTQFEDCTVLTIAHRLNTIMDYTRVI 1498
DB 1439 CLARALLRKTKILVLDATAVLETDLDLQISTRTTQFEDCTVLTIAHRLNTIMDYTRVI 1498
QY 1499 VLDKGEIQEYGAPELDLQORGLFYSMADAGLV 1531
DB 1499 VLDKGEIQEYGAPELDLQORGLFYSMADAGLV 1531
QY 1531 VLDKGEIQEYGAPELDLQORGLFYSMADAGLV 1531
DB 1531 VLDKGEIQEYGAPELDLQORGLFYSMADAGLV 1531

RESULT 11
Q810E4 ID Q810E4 PRELIMINARY; PRT; 1532 AA.
AC Q810E4;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ATP-binding cassette protein C1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OC NCB TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Spleen;
RA Yabuuchi H., Takayanagi S., Ishikawa T.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AF487549; AA085437.1; -.
DR HSSP; P08716; 1MT0.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.

```

DR GO: 0042626; F:ATPase activity, coupled to transmembrane m. . . ; IEA.  
 DR GO: 0000166; F:nucleotide binding; IEA.  
 DR GO: 0005215; F:transporter activity; IEA.  
 DR GO: 0006810; P:transport; IEA.  
 DR InterPro: IPR003593; AAA\_ATPase.  
 DR InterPro: IPR011527; ABC\_membrane\_1.  
 DR InterPro: IPR001140; ABC\_TM\_transp.  
 DR InterPro: IPR003439; ABC\_transporter.  
 DR InterPro: IPR001395; Aldo/ket\_red.  
 DR InterPro: IPR005292; MRP\_assoc.  
 DR Pfam: PF00664; ABC\_membrane\_2.  
 DR Pfam: PF00005; ABC\_tran; 2.  
 DR ProDom: PD000006; ABC\_transporter; 2.  
 DR SMART: SM00382; AAA; 2.  
 DR TIGRFAM: TIGR00957; MRP\_assoc\_pro; 1.  
 DR PROSITE: PS00299; ABC\_TM1F; 2.  
 DR PROSITE: PS00211; ABC\_TRANSPORTER\_1; 2.  
 DR PROSITE: PS00893; ABC\_TRANSPORTER\_2; 2.  
 DR PROSITE: PS00063; ALDO\_KETO\_REDUCTASE\_3; UNKNOWN\_1.  
 KW ATP-binding.  
 SQ SEQUENCE 1532 AA; 171491 MW; 286939F63F5A3F68 CRC64;  
 Query Match 69.2%; Score 6932.5; DB 2; Length 1532;  
 Best Local Similarity 87.3%; Pred. No. 0;  
 Matches 1339; Conservative 99; Mismatches 92; Indels 3; Gaps 3;  
 QY 1 MALRGFCSADGSDPLWDMVNTWNTSNPDKFQNTLVWVPCFYLMACFPFYLYLSRH 60  
 DB 1 MALSFCSGSDGSDPLWDMVNTWNTSNPDKFQNTLVWVPCFYLMACFPFYLYLSRH 60  
 QY 61 DRGYTQMTPLNKTALGFLMIVCWADLFYFWSRSGIFLAPVFLSPPTLLGITLLA 120  
 DB 61 DRGYTQMTPLNKTALGFLMIVCWADLFYFWSRSGIFLAPVFLSPPTLLGITLLA 120  
 QY 121 TFLIOLERRKGQSSGIMTLFWALVLCALATLRSKIMTALKEDAQVDLFRDITFVYFS 180  
 DB 121 TFLIOLERRKGQSSGIMTLFWALVLCALATLRSKIMTALKEDAQVDLFRDITFVYFS 180  
 QY 181 LLLIOLVLSFCSDRSPFSETIHDNCPPESSASFLSRTFWITGLIVRGYRQPLEGSD 240  
 DB 181 LVFIOLVLSFCSDRSPFSETIHDNCPPESSASFLSRTFWITGLIVRGYRQPLEGSD 240  
 QY 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPVVYS-SKDPAQPKSSKVDANEVEEA 299  
 DB 241 LWSLNKEDTSEVVPVLVNNWKECVKSRQVRIYAPPKDTPKPGSSQLDVNEVEEA 300  
 QY 300 LTVKSPQKEWNPFLKVLKTKTPGPFMSFFKALHLMFSGPOILKLIKFNVDTKAP 359  
 DB 301 LTVKSHKDRDPSFLKVLKTKTPGPFMSFLYKALHLMFAGPELELINFVNDREAP 360  
 QY 360 DMQGYFYTVLLFVTAQLTLVHQYFHCIFVSGMRIKTAIVGAVYRKALVITNSARKSST 419  
 DB 361 DMQGYLYTALLFVSACLQTLALHQYFHCIFVSGMRIKTAIVGAVYRKALVITNSARKSST 420  
 QY 420 VGEIVNLMSVDAQRFMDLATYINMTWASAPLOVLALYLWNLGSPSLAGVAVMLMPVP 479  
 DB 421 VGEIVNLMSVDAQRFMDLATYINMTWASAPLOVLALYLWNLGSPSLAGVAVMLMPVP 480  
 QY 480 NAVMAWTKTYQVAHMKSKDNRIKLMEITLNGIKVLKLYAWELAFKDKVLATROBELKVL 539  
 DB 481 NAVMAWTKTYQVAHMKSKDNRIKLMEITLNGIKVLKLYAWELAFKDKVNNIRQBELKVL 540  
 QY 540 KKSAYLSAVGTFTWCTPPLVALCTFAVYVTTIDENNILDNAOTAFVSLALFNILRPLNLL 599  
 DB 541 KKSAYLSAVGTFTWCTPPLVALCTFAVYVTTIDENNILDNAOTAFVSLALFNILRPLNLL 600  
 QY 600 PWISSIVQASVSLKRLRIFLSHELEPPDSIERPVPKGGTNSITVRNATTTWARSPP 659  
 DB 601 PWISSIVQASVSLKRLRIFLSHELEPPDSIERWSIKDGGGNSITVRNATTTWARSPP 660  
 QY 660 TLNGITTFEPGALVAVGVQCGKSSLLSALLAEMDKVEGHVAKGSAVYVPOQAWION 719  
 DB 661 TLNGITTFEPGALVAVGVQCGKSSLLSALLAEMDKVEGHVAKGSAVYVPOQAWION 720

QY 720 DSLRENILFGCOLEBPYRSYVIOACALLPDLBILPSGDRTEIGEGVNLSCGQKRVSLA 779  
 DB 721 DSLRENILFGRLQEHYKAVMEACALLPDLBILPSGDLTEIGEGVNLSCGQKRVSLA 780  
 QY 780 RAVYNSADIYLPDDPLSAVDAHVGKHIPENVIPGPKMLKNKTRILVTHSKSVLPQVDVVI 839  
 DB 781 RAVYNSDIYLLDDPLSAVDAHVGKHIPEKVGPMGLLKNKTRILVTHGISLVPQVDVVI 840  
 QY 840 VMSGGKISEMGSYQBELLDAGAFAPLRTYASTEQDAEENGVTGVSFGKEAKOMENG 899  
 DB 841 VMSGGKISEMGSYQBELLDAGAFAPLRTYASTEQDAEENGVTGVSFGKEAKOMENG 900  
 QY 900 MLVDSAGKQQLQRLSSSSSYSGDISRHNSIAEIQAKAEKKEETWKLMEADKAQTQGVK 959  
 DB 901 ILVTDVAGKPLQRLHLSNSSSHSVTNQSHSIAELQKS-GVKEETWKLMEADKAQTQGVK 959  
 QY 960 LSVYNDYKATGLFISLIFLPMCHVYSALASNTWLSLWTD-PVNGTQOEHKTVLSV 1018  
 DB 960 LSVYNYNYKATGLFISLIFLPMCHVYSALASNTWLSLWTD-PVNGTQOEHKTVLSV 1019  
 QY 1019 YGALGISQIAVFGYSMAVSIIGILASRCLHVDLHLSILRSPMSFFERTPSGNLNVRFSK 1078  
 DB 1020 YGALGILQVAVFGYSMAVSIIGIFASRLHLDLQNLVLRSPMSFFERTPSGNLNVRFSK 1079  
 QY 1079 ELDTVDSMIPVIMKFMGSLFNVIACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSR 1138  
 DB 1080 ELDTVDSMIPVIMKFMGSLFNVIACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSR 1139  
 QY 1139 OLKRLSVSRSPVSHFNELTGLSVIRAFEOERFIHQSDLKVDENOKAYVPSIVANRW 1198  
 DB 1140 OLKRLSVSRSPVSHFNELTGLSVIRAFEOERFIHQSDLKVDENOKAYVPSIVANRW 1199  
 QY 1199 LAVRLVCVNCIVLFAALPAVIRSHLSAGLVGLSVYSLOVTVTLNMLVRMSSEMETNI 1258  
 DB 1200 LAVRLVCVNCIVLFAALPAVIRSHLSAGLVGLSVYSLOVTVTLNMLVRMSSEMETNI 1259  
 QY 1259 VAVERLKEYSETEKAPWQIQTAPPSWMPQVGRVFRNYCLRYREDLDFVLRLHNVITN 1318  
 DB 1260 VAVERLKEYSETEKASWQIQTAPPSWMPQVGRVFRNYCLRYREDLDFVLRLHNVITN 1319  
 QY 1319 GGEKGVIGRTGAGKSSITGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPV 1378  
 DB 1320 GGEKGVIGRTGAGKSSITGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPV 1379  
 QY 1379 LPSGSLRNWLPFSSYSDVEEVTSLAHLKDFVSALPKLDPHECAEGENLSVGORQLV 1438  
 DB 1380 LPSGSLRNWLPFSSYSDVEEVTSLAHLKDFVSALPKLDPHECAEGENLSVGORQLV 1439  
 QY 1439 CLARALLRKTILVLDATAAVDLETDLLIOSTIRTFQEDCTVLTIAHRLNTIMDYTRVI 1498  
 DB 1440 CLARALLRKTILVLDATAAVDLETDLLIOSTIRTFQEDCTVLTIAHRLNTIMDYTRVI 1499  
 QY 1499 VLDKGEIYEGNAPSLLQORGLFYSMKADAGLV 1531  
 DB 1500 VLDKGEIYEGNAPSLLQORGLFYSMKADAGLV 1532

RESULT 12

Q8CG09 PRELIMINARY; PRT; 1532 AA.  
 ID Q8CG09  
 AC Q8CG09;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Multidrug resistance-associated protein 1.  
 GN Name=Mp1;  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN=Sprague-Dawley; TISSUE=Brain;  
 RA Yang Z., Li C.S.W., Shen D.D., Ho R.J.Y.;  
 RL Submitted (Oct-2002) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: Belongs to the ABC transporter family.  
 DR EMBL; AY170916; AA086532.1; -;  
 DR HSP; P08716; IMT0.  
 DR GO; GO:0016021; C: integral to membrane; IEA.  
 DR GO; GO:0005524; F: ATP binding; IEA.  
 DR GO; GO:0042626; F: ATPase activity, coupled to transmembrane m. . .; IEA.  
 DR GO; GO:0001666; F: nucleotide binding; IEA.  
 DR GO; GO:0005215; F: transporter activity; IEA.  
 DR GO; GO:0006810; P: transporter; IEA.  
 DR InterPro; IPR003593; AAA\_ATPase.  
 DR InterPro; IPR011527; ABC\_membrane\_1.  
 DR InterPro; IPR01140; ABC\_TM\_transp.  
 DR InterPro; IPR003439; ABC\_transporter.  
 DR InterPro; IPR001395; Aldo\_ket\_red.  
 DR InterPro; IPR005292; MRP\_assoc.  
 DR Pfam; PF00664; ABC\_membrane; 2.  
 DR Pfam; PF00005; ABC\_tran; 2.  
 DR ProDom; PD000006; ABC\_transporter; 2.  
 DR SMART; SM00382; AAA; 2.  
 DR TIGRfam; TIGR00957; MRP\_assoc\_pro; 1.  
 DR PROSITE; PS00929; ABC\_TMIF; 2.  
 DR PROSITE; PS0211; ABC\_TRANSPORTER\_1; 2.  
 DR PROSITE; PS0893; ABC\_TRANSPORTER\_2; 2.  
 DR PROSITE; PS00063; ALDO\_KETO\_REDUCTASE\_3; UNKNOWN\_1.  
 KW ATP-binding.  
 SQ SEQUENCE 1532 AA; 171557 MW; C18P3554DD85732A CRC64;

Query Match 69.2%; Score 6932.5; DB 2; Length 1532;  
 Best Local Similarity 87.3%; Pred. No. 0;  
 Matches 1338; Conservative 100; Mismatches 92; Indels 3; Gaps 3;

QY 1 MALRGFCSADGSDPLDWDNVTWNTNPDFTKCFQNTLVWVWVFCFYLWACFPFYLYLSRH 60  
 DB 1 MALRFCSGSDGSDPLDWDNVTWNTNPDFTKCFQNTLVWVWVFCFYLWACFPFYLYLSRH 60

QY 61 DRGYIQMTPLNKTKTALGFLMIVCWADLFYFWSRSGIFLAPVFLVPSLTLLGTTLLA 120  
 DB 61 DRGYIQMTHLNKAKTALGFLMIICWADLFYFWSRSGQMLLAPVLLVPSLTLLGTTMLLA 120

QY 121 TELIQLERRKGVOSSGIMLTFWLVVALCALILRSKIMTALKEDAQVDLFRDITFYVYFS 180  
 DB 121 TELIQFERRKGVOSSGIMLTFWLVVALCALILRSKIIISALKDAQVDMFRDSAPLYFT 180

QY 181 LILLIQLVLSFSDRPLFSETIHDNCPPESSASFLSRITFWITGLIVRGYRQPLEGSD 240  
 DB 181 LVFIQVLVLSFSDSPLFSEVTRDNPCESSASFLSRITFWITGMMVQGYRQPLKSSD 240

QY 241 LWSLNKEDTSQGVPLVKNWKKCECAKTRKQPKVYVYS - SKDPAQPKESSKVDANEVEA 299  
 DB 241 LWSLNKEDTSBEVPLVNNWKKCECVKSRKQPVRIYVAPPDPTKPKGSSQLDYNEVEA 300

QY 300 LIVKSPQKEMPSLFKVLKYTFGRFELMSFFKAHDLMMSGQILKLLIKFVNDYKAP 359  
 DB 301 LIVKSHKDRDPSLFKVLKYTFGRFELMSFLYKALHDLMMFAGFEILELIINFVNDREAP 360

QY 360 DWQGYFTVLLFVTAQLTLVHQYFHCIFYSGMRITKAVIGAVYRKALVITNSARKSST 419  
 DB 361 DWQGYLTALLFVSACLTQALHQYFHCIFYTGMRIKTAUVGVYRKALVITNSARKSST 420

QY 420 VGEIVNLSVDAQRFMDLATINNIWSAPLOVILALYLLMLNLGFSVLAVGVNMLVMPV 479  
 DB 421 VGEIVNLSVDAQRFMDLATINNIWSAPLOVTLALYFLMLNLGFSVLAVGVNMLVMPV 480

QY 480 NAVMAMTKTYQVAHMKSKONRIKLMNEILNGIKVLKLYAWELAFKDLALIOBELKVL 539  
 DB 481 NAVMAMTKTYQVAHMKSKONRIKLMNEILNGIKVLKLYAWELAFQDKVMNIOEELKVL 540

QY 540 KKSAYLSAAGTFTVCTPFLVALCTFAVYVITDENNILDQAOTAFVSLALFNILRFPNLL 599  
 DB 541 KKSAYLAAGVTFTVCTPFLVALSTFAVFTVDEKNILDAKAFVSLALFNILRFPNLL 600

QY 600 PMVISSIVQASVSLKRLIFLSHEELEPDSIERRPVKCGGGTNSITVRNATFTWARSDDP 659  
 DB 601 PMVISSIVQASVSLKRLIFLSHEELEPDSIERRSVKCGGGTNSITVRNATFTWARSDDP 660

QY 660 TLNGITFSPGALVAVVQVCGGKSSLLSALLAEMDKVEGHVAKGSVAYVPPQAWIQN 719  
 DB 661 TLNGITFAIPDGLVAVVQVCGGKSSLLSALLAEMDKVEGHVAKGSVAYVPPQAWIQN 720

QY 720 DLSRENILFGCOLPEPYRIVIOACALLPDLILPSGDRTEIGEGVNLGGCKORVSLA 779  
 DB 721 DLSRENILFGRPLQCHCYKAVWEACALLPDLILPSGDLTEIGEGVNLGGCKORVSLA 780

QY 780 RAVYNSADIYLFDDPLSAVDHVGKHIFENVIKPKMLKKNKTRILVTHSMYSLPQVDVII 839  
 DB 781 RAVYNSDIYLLDDPLSAVDHVGKHIFEKVVGPMLLKNKTRILVTHGISYLPQVDVII 840

QY 840 VMSGGKI SEMGSYOELLARDGAFALFYASTEQEODAEENGVTGVGPGKEAKOMENG 899  
 DB 841 VMSGGKI SEMGSYOELLARDGAFALFYASTEQEODAEENGVTGVGPGKEAKOMENG 900

QY 900 MLVTD SAGKQLOQI SSSSSYSYSGDISRHHNSTAELOKAEKKEETWKLMEADKATGVVK 959  
 DB 901 ILVTD AVGKPLQRLHLSNSSSHSVTVNQHSSTAELOKS -GVKEETWKLMEADKATGVVK 959

QY 960 LSVYWDYKAI GLFISFLSIFLFCMCHVSAALSNYWSLWTD -PIVNGTQEHKTVRLSV 1018  
 DB 960 LSVYWNMYKAI GLCISFLSIFLFCMCHVSAALSNYWSLWTD -PIVNGTQEHKTVRLSV 1019

QY 1019 YGALGISOGIAVFGYSMAVSI GGILASRLCHVDLHLSILRSPMFFERTSGNLVNRSK 1078  
 DB 1020 YGALGILQVAVFGYSMAVSI GGIFASRLHLDLQNLVLRSPMFFERTSGNLVNRSK 1079

QY 1079 ELDTVDSMIPKVIKMFMSLENFVIGACIVILLATPIAAIIIPPLGLIYFFVQRFVASSR 1138  
 DB 1080 ELDTVDSMIPQVIKMFMSLSFVIGAVIILLATPIAAVILPPLGLVYFFVQRFVASSR 1139

QY 1139 QLKRLSVSRPVYSHNETLLGVSVIRAFEEORFTHOSDLKVDENQKAYPSIVANRW 1198  
 DB 1140 QLKRLSVSRPVYSHNETLLGVSVIRAFEEORFIRQSDLKVDENQKAYPSIVANRW 1199

QY 1199 LAVRLECVGNCIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTYLNLVMSSEMETNI 1258  
 DB 1200 LAVRLECVGNCIVLFAALFAVISRHSLSAGLVGLSVSYSLQITAYLNLVMSSEMETNI 1259

QY 1259 VAVERLKEYSETEKAPWQIETAPPSSWPQGVFEFRNYCLRYREDLDVFLRHINVTIN 1318  
 DB 1260 VAVERLKEYSETEKASWQIETAPPSTWPHSGRVEFRDYCLRYREDLDVFLKHINVTIE 1319

QY 1319 GGEKVGIVGRTGAKSSLTGLFRINESAGEIIGDGINAKTGLHDLRKITIIPODPV 1378  
 DB 1320 GGEKVGIVGRTGAKSSLTGLFRINESAGEIIGDGINAKTGLHDLRKITIIPODPV 1379

QY 1379 LFGSLRMNLDPPFSQYSDSEEVWTSLELAHLKDFVSALPDKLDHECAGGENLSVGQQLV 1438  
 DB 1380 LFGSLRMNLDPPFSQYSDSEEVWMALELAHLKGFVSALPDKLNHECAGGENLSVGQQLV 1439

QY 1439 CLARALLRKTKIILVDEATAVDETDLLQSIQIRTOFEDCTVLTIAHRLNTIMDYTRVI 1498  
 DB 1440 CLARALLRKTKIILVDEATAVDETDLLQSIQIRTOFEDCTVLTIAHRLNTIMDYTRVI 1499

QY 1499 VLDKGEIQEYGAPELQQRGLFYSMAKQAGLV 1531  
 DB 1500 VLDKGEIREGAPSELLQQRGLFYSMAKQAGLV 1532

RESULT 13  
 Q810G9 ID Q810G9 PRELIMINARY; PRT; 1523 AA.  
 AC Q810G9;  
 DT 01-JUN-2003 (T-EMBLrel. 24, Created)  
 DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)  
 DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)

DE ATP-binding cassette protein C1 variant A.  
GN Name=Abcc1;  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_Taxid=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Sprague-Dawley; TISSUE=Spleen;  
RA Takayanagi S., Iehikawa T.;  
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: Belongs to the ABC transporter family.  
DR EMBL; AY174892; AAO44983.1; -;  
DR HSSP; P08716; 1MTO.  
DR GO; GO:0016021; C: integral to membrane; IEA.  
DR GO; GO:0005524; P: ATP binding; IEA.  
DR GO; GO:0042626; P: ATPase activity, coupled to transmembrane m. . .; IEA.  
DR GO; GO:0000166; F: nucleotide binding; IEA.  
DR GO; GO:0005215; F: transporter activity; IEA.  
DR GO; GO:0006810; P: transport; IEA.  
DR InterPro; IPR0013593; AAA ATPase.  
DR InterPro; IPR011527; ABC membrane 1.  
DR InterPro; IPR001140; ABC TM transp.  
DR InterPro; IPR003439; ABC transporter.  
DR InterPro; IPR001395; ABC transporter.  
DR InterPro; IPR005292; MRP assoc.  
DR Pfam; PF00664; ABC membrane 2.  
DR Pfam; PF00005; ABC tran; 2.  
DR ProDom; PD000006; ABC transporter; 2.  
DR SMART; SM00382; AAA; 2.  
DR TIGRFam; TIGR00957; MRP assoc. pro; 1.  
DR PROSITE; PS00929; ABC TM1F; 2.  
DR PROSITE; PS00211; ABC TRANSPORTER 1; 2.  
DR PROSITE; PS00893; ABC TRANSPORTER 2; 2.  
DR PROSITE; PS00063; ALDOXETO\_REDUCTASE\_3; UNKNOWN\_1.  
KW ATP-binding.  
SQ SEQUENCE 1523 AA; 170503 MW; E40337051A1CB9C6 CRC64;

Query Match 68.8%; Score 6892; DB 2; Length 1523;  
Best Local Similarity 87.0%; Pred. No. 0;  
Matches 1333; Conservative 98; Mismatches 90; Indels 12; Gaps 4;

QY 1 MALRGCSDGSDPLMDNWTNTNPDFTKCFQNTLVVWPCFYLWACFPFYLVSRLH 60  
DB 1 MALRGCSDGSDPLMDNWTNTHSTNPDFTKCFQNTLVVWPCFYLWACFPFYLVSRLH 60  
QY 61 DRGYQMTPLNKTALGFLMLVWADLFYFWSRSRGIPLAPFLVSPITLLGTTLLA 120  
DB 61 DRGYQMTPLNKTALGFLMLVWADLFYFWSRSRQGMLLAPVLLVSPITLLGTTLLA 120  
QY 121 TFLIOLERRKGVOSSGIMLTFLWVALVCLALILRSKIMTALKEDAQVDLFDITFVYVFS 180  
DB 121 TFLIOLERRKGVOSSGIMLTFLWVALVCLALILRSKIIISALKKDAQVDLFDITFVYVFS 180  
QY 181 LLLIOLVLCFSDRSPFSETHDNPCCPSSASFLSRITFWITGLIVRGYRQPLEGSD 240  
DB 181 LVFIQLVLCFSDRSPFSETHDNPCCPSSASFLSRITFWITGLIVRGYRQPLEGSD 240  
QY 241 LWSLNKEDTSEQVPLVQWKKCKAKTRQPKVYVS-SKDPAPQKSSKYDANEVEA 299  
DB 241 LWSLNKEDTSEVPLVNNWKKCKVSRQPVRIYVAPPDPTKPKGSSQLDVAEVEA 300  
QY 300 LLVKSPQKWNPSLKFVLYKTCFPLMSFFFKAIHDLMMFSGPQLKLLIKFVNDTKAP 359  
DB 301 LLVKSHKDRDPSLFVLYKTCFPLMSFLYKALHDLMMFAGPELLELIINFVNDREAP 360  
QY 360 DMQGYFYTVLLFVYACQLTLVHQYFHCIFVSGMRKTKAVIGAVYRKALVITNSARKST 419  
DB 361 DMQGYLYTALLFVSACQLTLVHQYFHCIFVTGMRKTKAVIGAVYRKALVITNSARKST 420  
QY 420 VGEIVNLSVDAQRFMDLATYINMIWSAPLOVTLALYLWNLNLPSPVLGAVVWLMVVP 479  
DB 421 VGEIVNLSVDAQRFMDLATYINMIWSAPLOVTLALYLWNLNLPSPVLGAVVWLMVVP 480

QY 480 NAVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLYAWELAFKDKVLAIRQBELKYL 539  
DB 481 NAVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLYAWELAFKDKVLMNIRQBELKYL 540  
QY 540 KKSAYLSAVGPTTWCTPFLVALCTFAVYVTTDENWILDQAQTAFLVSLALFNILRPPLNL 599  
DB 541 KKSAYLSAVGPTTWCTPFLVALCTFAVYVTTDENWILDQAQTAFLVSLALFNILRPPLNL 600  
QY 600 PMVISSIVQASVSLKRLAIFLSHEELPDSIERRPVKGGGNSITVRNATPTWARSDEP 659  
DB 601 PMVISSIVQASVSLKRLAIFLSHEELPDSIERRPVKGGGNSITVRNATPTWARSDEP 660  
QY 660 TLNGITFISPEGALVAVVQVCGCKSSLSALLAEMDKVEGHVAIKGSVAYVPPQAWIQN 719  
DB 661 TLNGITFISPEGALVAVVQVCGCKSSLSALLAEMDKVEGHVAIKGSVAYVPPQAWIQN 720  
QY 720 DSLRENILFGCOLLEPYRSVLOACALLPDLBILPSGDRTEIGEGKVNLSGQKQVSLA 779  
DB 721 DSLRENILFGCOLLEPYRSVLOACALLPDLBILPSGDRTEIGEGKVNLSGQKQVSLA 780  
QY 780 RAVYSNADILYLPDDPLSAVDHVGKHIPENVIGPKGMLKNKTRILVTHSMSYLPQVDVIL 839  
DB 781 RAVYSNADILYLPDDPLSAVDHVGKHIPENVIGPKGMLKNKTRILVTHSMSYLPQVDVIL 840  
QY 840 VMSGGKISEMGSYQBELLDGAFABFLRTYASTEQEQDAENGVTGVSFGKEAKQWENG 899  
DB 841 VMSGGKISEMGSYQBELLDGAFABFLRTYASTEQEQDAENGVTGVSFGKEAKQWENG 900  
QY 900 MLVTSAGKQRLQSSSSSYGDISRHNSHTAEIQKAEKKEETWKLMEADKAQGVQVK 959  
DB 901 MLVTSAGKQRLQSSSSSYGDISRHNSHTAEIQKAEKKEETWKLMEADKAQGVQVK 960  
QY 960 LSVYNDYMKACIGLFTSIFLPMCNHVSALASNTWLSLWTD-PVNGTQEHKTVRLSV 1018  
DB 951 LSVYNDYMKACIGLFTSIFLPMCNHVSALASNTWLSLWTD-PVNGTQEHKTVRLSV 1019  
QY 1019 YGALGISQIAVFGYSMAVSIIGGILASRCLHVDLLHSILRSPMSFPFRTSGNLNRFSK 1078  
DB 1011 YGALGISQIAVFGYSMAVSIIGGILASRCLHVDLLHSILRSPMSFPFRTSGNLNRFSK 1079  
QY 1079 ELDTVDSMIPVVKPMGSLFNVIACIVILLATPIAIIIPPLGLIYFFVQRYFVASSR 1138  
DB 1071 ELDTVDSMIPVVKPMGSLFNVIACIVILLATPIAIIIPPLGLIYFFVQRYFVASSR 1139  
QY 1139 QKRLSVRSRSPVYSHFNETLLGVSVIRAFEOERFIHOSDLKVDENQKAYPSIVANRW 1198  
DB 1131 QKRLSVRSRSPVYSHFNETLLGVSVIRAFEOERFIHOSDLKVDENQKAYPSIVANRW 1199  
QY 1199 LAVRLCEVGNCTVLPFAALFAVISHRSLSAGLVGLSVSYSLQITAYLVNMLVRMSSEMETNI 1258  
DB 1191 LAVRLCEVGNCTVLPFAALFAVISHRSLSAGLVGLSVSYSLQITAYLVNMLVRMSSEMETNI 1259  
QY 1259 VAVERLKEYSETKEAPQIQTETAPPSWPQVGRVFRNYCLRYREDLDFVLRHINVTIN 1318  
DB 1251 VAVERLKEYSETKEAPQIQTETAPPSWPQVGRVFRNYCLRYREDLDFVLRHINVTIN 1319  
QY 1319 GGEKGVIVGRTGAGKSLTLGLFRINESAEGIIIDGINAKIHLHDLFKFTIIPQDPV 1378  
DB 1311 GGEKGVIVGRTGAGKSLTLGLFRINESAEGIIIDGINAKIHLHDLFKFTIIPQDPV 1379  
QY 1379 LFSGSLRWMLDPPFSQYSDVEEVMVALEHLKGFVSALPDKLNHECAEGGENISVGQRLV 1438  
DB 1371 LFSGSLRWMLDPPFSQYSDVEEVMVALEHLKGFVSALPDKLNHECAEGGENISVGQRLV 1439  
QY 1439 CLARALLKTKILVDEATAAVDLETDLLIQSTIRTFQEDCTVLTIAHLNLTIMDYTRI 1498  
DB 1431 CLARALLKTKILVDEATAAVDLETDLLIQSTIRTFQEDCTVLTIAHLNLTIMDYTRI 1499  
QY 1499 VLDKGEIYEGAPSDLLQOGLFYSMKADAGLV 1531  
DB 1491 VLDKGEIYEGAPSDLLQOGLFYSMKADAGLV 1532

## RESULT 14

Q68CP7 PRELIMINARY; PRT; 1215 AA.  
 AC Q68CP7; 25-OCT-2004 (T-EMBLrel. 28, Created)  
 DT 25-OCT-2004 (T-EMBLrel. 28, Last sequence update)  
 DT 25-OCT-2004 (T-EMBLrel. 28, Last annotation update)  
 DE Hypothetical protein DKF2p781G125.  
 GN Name=DKF2p781G125;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Amalgam;  
 RG The German cDNA Consortium;  
 RA Koehrer K., Beyer A., Mewes H.W., Weil B., Amid C., Oeanger A.,  
 RA Fobo G., Han M., Wiemann S.;  
 RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; CR749835; CAH18691.1; -  
 DR InterPro; IPR003593; AAA\_ATPase.  
 DR InterPro; IPR001140; ABC\_TM\_transp.  
 DR InterPro; IPR003439; ABC\_transporter.  
 DR InterPro; IPR005292; MRP\_assoc.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR Pfam; PF00664; ABC\_membrane; 2.  
 DR Pfam; PF00005; ABC\_tran; 2.  
 DR ProDom; PD000006; ABC\_transporter; 2.  
 DR SMART; SM00382; AAA; 2.  
 DR TIGRFAMs; TIGR00957; MRP\_assoc\_pro; 1.  
 DR PROSITE; PS00211; ABC\_TRANSPORTER\_1; 2.  
 DR PROSITE; PS00893; ABC\_TRANSPORTER\_2; 2.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; UNKNOWN\_1.  
 KW ATP-binding; Hypothetical protein.  
 SQ SEQUENCE 1215 AA; 134926 MW; E490ADIP5268F72E CRC64;

Query Match 60.7%; Score 6075; DB 2; Length 1215;  
 Best Local Similarity 98.9%; Pred. No. 0;  
 Matches 1202; Conservative 0; Mismatches 3; Indels 10; Gaps 1;

QY 327 MSFFPKAHLDMFSGPQILKLIKFNVDNTPADMQGYFYVLLFVTTACQLTVLHQQYFH 386  
 DB 1 MSFFPKAHLDMFSGPQILKLIKFNVDNTPADMQGYFYVLLFVTTACQLTVLHQQYFH 60

QY 387 ICFVSGMIKTAVIGAVYRKALVITNSARKSTVGEIVNLSVDAQRFMDLATYINNIWS 446  
 DB 61 ICFVSGMIKTAVIGAVYRKALVITNSARKSTVGEIVNLSVDAQRFMDLATYINNIWS 120

QY 447 APLQVILALYLLNLNLPSPVLGAVVMVLPVNAVAMKTKTYQVAHMKSKDNRIKLMN 506  
 DB 121 APLQVILALYLLNLNLPSPVLGAVVMVLPVNAVAMKTKTYQVAHMKSKDNRIKLMN 180

QY 507 EILNGIKVLKYANELAFKDKVLAIRQBELKVLKKSAYLSAVGTFTWCTPFLVALCTFA 566  
 DB 181 EILNGIKVLKYANELAFKDKVLAIRQBELKVLKKSAYLSAVGTFTWCTPFLVALCTFA 240

QY 567 VYVTIDENNILDAQTAFLVSLALFNILPPLNIPWISSIVQASVSLKRLRIFLSHEELE 626  
 DB 241 VYVTIDENNILDAQTAFLVSLALFNILPPLNIPWISSIVQASVSLKRLRIFLSHEELE 300

QY 627 PDSITERRPVKGGTNSITVNRNATFTWASDPPTLNGITPSIPGALVAVVQVCGCKSS 686  
 DB 301 PDSITERRPVKGGTNSITVNRNATFTWASDPPTLNGITPSIPGALVAVVQVCGCKSS 360

QY 687 LLSALLAEMDKVEGHVAIKGSVAVVPOQAWTQNDLSRENILFGQLEEPYRYSVIQACAL 746  
 DB 361 LLSALLAEMDKVEGHVAIKGSVAVVPOQAWTQNDLSRENILFGQLEEPYRYSVIQACAL 420

QY 747 LPDLEILPSGDRTEIGKGNVLSGGQKQKRVSLAPVSNADYILFDDPPLSAVDHVGKHI 806  
 DB 421 LPDLEILPSGDRTEIGKGNVLSGGQKQKRVSLAPVSNADYILFDDPPLSAVDHVGKHI 480

QY 807 FENVIGPKGMLKNKTRILVTHSMSYLPQVDVIIVMSGGKISEMGSYQELLARDGAFAPFL 866  
 DB 481 FENVIGPKGMLKNKTRILVTHSMSYLPQVDVIIVMSGGKISEMGSYQELLARDGAFAPFL 540

QY 867 RYASTEQEQDABEN-----GVTGVSQPGKEAKQEMNGMLVTDTSAGKOLQRLSS 916  
 DB 541 RYASTEQEQDABEN-----GVTGVSQPGKEAKQEMNGMLVTDTSAGKOLQRLSS 600

QY 917 SSSYSGDISRHHNSTAEQKAEETKWEADKATGOVKLSVYNDYKKAIGLFTSF 976  
 DB 601 SSSYSGDISRHHNSTAEQKAEETKWEADKATGOVKLSVYNDYKKAIGLFTSF 660

QY 977 LSIFLFCMNHVSALASNYLWLSLWTDPIVNGTQEHTKVRLSVYGAIGISQGIATVGYSWA 1036  
 DB 661 LSIFLFCMNHVSALASNYLWLSLWTDPIVNGTQEHTKVRLSVYGAIGISQGIATVGYSWA 720

QY 1037 VSIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVRNRFKELDTVDSMTPEVIKMFNG 1096  
 DB 721 VSIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVRNRFKELDTVDSMTPEVIKMFNG 780

QY 1097 SLFNVIGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQLKRLSVSRSPVYSHFN 1156  
 DB 781 SLFNVIGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQLKRLSVSRSPVYSHFN 840

QY 1157 ETLIGVSVIRAFEEQERFIHOSDLKVDENQKAYYPSIVANRWLAVRLECVGNCIVLFAAL 1216  
 DB 841 ETLIGVSVIRAFEEQERFIHOSDLKVDENQKAYYPSIVANRWLAVRLECVGNCIVLFAAL 900

QY 1217 FAVISRHSLSAGLVGLSVSYSLQVTVYTNLVRMSMETNIVAVERLKEYSETEKEAPW 1276  
 DB 901 FAVISRHSLSAGLVGLSVSYSLQVTVYTNLVRMSMETNIVAVERLKEYSETEKEAPW 960

QY 1277 QIQTAPPSSWPQVGRVFRNYCLRYREDLDFVLRHINVTINGGKVGIVGRTGAGKSSL 1336  
 DB 961 QIQTAPPSSWPQVGRVFRNYCLRYREDLDFVLRHINVTINGGKVGIVGRTGAGKSSL 1020

QY 1337 TLGLFRINESAGEIIIDGINIAKIGHDLRFKTIIPQDPVLFSGLRNLNLPFSQVSD 1396  
 DB 1021 TLGLFRINESAGEIIIDGINIAKIGHDLRFKTIIPQDPVLFSGLRNLNLPFSQVSD 1080

QY 1397 EEVWTSLELAHLKDFVSALPDKLDHECAEGENLSVGQROLVCLARALLRKTILVLDEA 1456  
 DB 1081 EEVWTSLELAHLKDFVSALPDKLDHECAEGENLSVGQROLVCLARALLRKTILVLDEA 1140

QY 1457 TAAVDLETDDLIOSTIRTFQEDCTVLTIAHLNLTIMDYTRVIVLDKGEIOEYGAPSDLLQ 1516  
 DB 1141 TAAVDLETDDLIOSTIRTFQEDCTVLTIAHLNLTIMDYTRVIVLDKGEIOEYGAPSDLLQ 1200

QY 1517 QRGLFYSMAKDAGLV 1531  
 DB 1201 QRGLFYSMAKDAGLV 1215

RESULT 15  
 Q80ZK8 PRELIMINARY; PRT; 1519 AA.  
 ID Q80ZK8  
 AC Q80ZK8  
 DT 01-JUN-2003 (T-EMBLrel. 24, Created)  
 DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)  
 DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)  
 DE Similar to ATP-binding cassette, sub-family C (CFTR/MRP), member 3 (Fragment).  
 DE (Fragment).  
 GN Name=Abccc3;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FVB/N; TISSUE=Liver;  
 RA Strausberg R.;  
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.

Search completed: March 18, 2005, 11:12:24  
Job time : 214.058 secs

**THIS PAGE BLANK (USPTO)**